Psychological Monographs

General and Applied

o. 517
961

Shepard, Hovland, and Jenkins

ol. 75
o.13

Learning and Memorization of Classifications

By

Roger N. Shepard
Bell Telephone Laboratories, Murray Hill, New Jersey

Carl I. Hovland
Yale University

and Herbert M. Jenkins
Bell Telephone Laboratories, Murray Hill, New Jersey

Price $2.00

Edited by Norman L. Munn

Published by the American Psychological Association, Inc.

Psychological Monographs:

General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology* with the *Psychological Monographs*

Manuscripts and correspondence on editorial matters should be sent to the Editor. *Psychological Monographs* publishes comprehensive experimental investigations and other psychological studies which do not lend themselves to adequate presentation as journal articles. Major space is given to the author's original contribution; introductory and bibliographic materials, as well as statistical tables and graphs, must be kept within reasonable bounds. Tables, graphs, and appendix materials which deal with detail not essential to adequate presentation of the findings may be made available through the American Documentation Institute—for details of this procedure, see the APA *Publication Manual*. Preparation of manuscripts for publication as monographs should follow the procedure given in the APA *Publication Manual*. Publication in *Psychological Monographs* is free of cost to the author, except in cases where early publication is requested or author's alterations are made in galley proofs.

Correspondence on business matters should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St., N.W., Washington 6, D.C. Address changes must arrive by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee third-class forwarding postage.

Vol. 75, No. 13 Whole No. 517, 1961

Psychological Monographs: General and Applied

LEARNING AND MEMORIZATION OF CLASSIFICATIONS[1]

ROGER N. SHEPARD
Bell Telephone Laboratories, Murray Hill, New Jersey

CARL I. HOVLAND
Yale University

AND HERBERT M. JENKINS
Bell Telephone Laboratories, Murray Hill, New Jersey

The present study explores some of the factors that determine how difficult a classification will be to learn or remember. By a "classification" we mean, here, simply a grouping of a given set of stimuli into two or more mutually exclusive and exhaustive classes. The learning or memorization of a classification can be regarded as a process of associating, to each stimulus, a certain response. This response might be the verbal label arbitrarily assigned to the class containing that stimulus, or it might be the act of sorting that stimulus into the bin arbitrarily assigned to its class. The essential feature of a *classification task,* however, is that the same response is assigned to several different stimuli. Accordingly, we reserve the term *identification task* for cases in which a different response is paired with each stimulus. In either case, the word "memorization" is intended, here, to refer to those conditions in which the materials to be learned are presented only once prior to the test for retention.

[1] Experiment I was carried out at the Bell Telephone Laboratories and Fairleigh Dickinson University. Experiments II and III were supported by grants from the Bell Telephone Laboratories and the Behavioral Sciences Division of the Ford Foundation. The support of both of these organizations is gratefully acknowledged. Thanks and deep appreciation are also expressed to the following: John Gibbon, for serving as experimenter in Experiment I; Earl Hunt, for conducting a pilot run for Experiment II and for assistance in setting up that experiment; Paul A. Lane and Samuel Squires of the Psychology Department of the University of Bridgeport, for permission to use their students as subjects; Ralph F. Garofalo and Ferdinand J. Fritzky, for serving as experimenters in Experiment II; and Albert Bregman, for serving as experimenter in Experiment III. Finally, we wish to thank W. J. McGill for his helpful comments concerning the manuscript.

Learning by Concept and Learning by Rote

In general, since the stimuli that are classified together need not be discriminated from each other, less information about a stimulus is required to classify it than to identify it. Therefore we might expect that classifications would be more easily learned and remembered than identifications. For example, if we have four horses and four dogs, we should certainly find it easier to remember one name for the horses and one for the dogs than to remember a different name for each of the eight individual animals. One is tempted to say that the difference, here, is between learning by concept and learning by rote. Horses presumably have something in common (not shared by the dogs) such that, after one name has been learned for three horses, the extension of this same name to the fourth horse requires little if any further learning. In the case of identification learning no such saving is possible. After a different name has been learned for each of three horses, the association of the fourth name to the fourth horse must still be formed *de novo,* i.e., by rote.

Unfortunately there are certain drawbacks to the use of a comparison between identification learning and classification learning for the purpose of clarifying the relation between rote and concept learning. First, as Bricker (1955) has pointed out, the reduction in the number of responses entailed by the conversion of the identification task into

a classification task also results in a change in the chance level of performance. For instance, with eight stimuli (as in our example), subjects (*S*s) who responded completely at random would on the average select the correct classifying response (out of the two alternatives) one-half of the time, but they would select the correct identifying response (out of the eight alternatives) only one-eighth of the time. Of course one could correct the obtained error scores for this difference in chance level or else use a different measure (such as trials to criterion) for which the difference in chance level might not be as great. But, even if we were to substantiate in this way that classifications are easier than identifications, we should still have to sort out the contributions of two different factors. For the reduction that would presumably be found in the difficulty of classification learning could be a consequence of (*a*) the fact that the stimuli that are classified together have some property in common with which the classificatory response can be associated (without distinguishing each stimulus from every other), or it could be a consequence of (*b*) the reduction simply in the *number* of responses that must be "kept in mind." Factor *a* seems central to concept learning, but *b* is presumably more akin to the length-of-list factor investigated in studies of rote learning.

Actually, the difficulty of a classification task can be changed radically without altering the set of stimuli or responses in any way but, rather, simply by modifying the assignment between them. Surely we should have more difficulty in learning one name for two of the horses and two of the dogs and the other name for the remaining two horses and dogs than (as in the example considered above) in learning one name for the horses and the other for the dogs. The cross-species classification evidently entails a larger component of rote learning and, might, indeed, be comparable in difficulty to learning a separate identifying response for each animal. Moreover the difference in difficulty of two such classifications could not be attributed either to changes in length of list or in chance expectation. Clearly, then, the extent to which the potential reduction in difficulty from identification to classification learning is realized depends upon how the stimuli are grouped together in their assignment to the responses. In particular, classification learning (with a fixed set of stimuli and responses) has been conclusively demonstrated to proceed more rapidly when the responses are assigned on the basis of common properties of the stimuli rather than in a completely arbitrary manner (French, 1953; Smith, 1954). And the same kind of result is found whether the stimuli are four-letter sequences (French, 1951), playing cards (Rogers, 1952), irregular closed curves (French, 1953), or other geometrical figures (Metzger, 1958; Smith, 1954; Wolfle, 1932). Metzger and Smith distinguish the two contrasting conditions of classification learning as "systematic" or "structured" concept tasks, on the one hand, and "random" concept tasks on the other. However, since the word "concept" seems to us to imply "systematic" or "structured," we prefer the somewhat more neutral word "classification" when conditions are included in which the stimuli are grouped by fiat. In any case, if the rote component of a classification task can be substantially changed simply by regrouping the stimuli in their assignment to the responses, some understanding of the relation between rote and concept learning might be gained by examining the performance of *S*s when the same set of stimuli is classified in different ways.

Characterization of Classifications in Terms of the Dimensions and Values of the Stimuli

In order to simplify the task of describing and controlling the properties of the stimuli we have confined our investigation to stimuli constructed by selecting one of two possible values on each of three different dimensions. For example, the dimensions might be size, color, and shape and the values on these might be large or small, black or white, and square or triangular. We then get the $2 \times 2 \times 2$ or eight geometrical figures shown in the box labeled I in Figure 1. These eight stimuli can then be classified in a very large number of ways. However, in order to

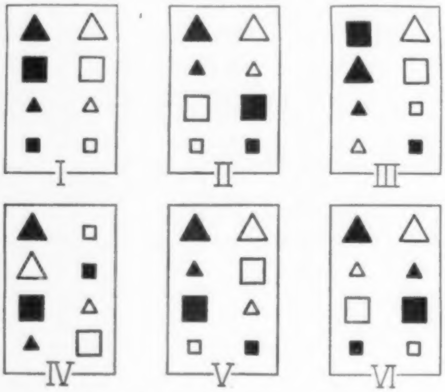

FIG. 1. Six different classifications of the same set of eight stimuli. (Within each box the four stimuli on the left belong in one class and the four stimuli on the right in the other class.)

equate the informational content of the different classifications (Hovland, 1952), we use only dichotomous classifications in which four of the eight stimuli are assigned to one response and the remaining four stimuli to the other.[2] The number of different classifications of this kind is given by the number of combinations of four things taken from eight: namely, $8!/(4!)^2$ or 70.

Six of these 70 possible classifications of the same eight stimuli are illustrated in the six boxes in Figure 1. In each case one response, say the letter A, might be assigned to the four stimuli on the left and another response, say B, to the four stimuli on the right. In Box I, then, the necessary and sufficient condition for the correct application of Response A is simply that the stimulus be *black*. The dimensions of size and shape are irrelevant for this classification. Another classification of these same stimuli that would presumably be somewhat more difficult to learn and remember is illustrated in Box II. Here the necessary and sufficient condition for Response A is that the stimulus be *either black and triangular or else white and square*. In this classification only the dimension of size is irrelevant. The Classifications III, IV, and V will be discussed later. The classification in Box VI, however, represents an extreme case and should therefore be considered now. The necessary and sufficient condition for Response A in this classification is that the stimulus be *either triangular, and large and black, or small and white; or else square, and large and white, or small and black*. Here, none of the three dimensions is irrelevant. If the difficulty of learning has any relation to the length of these rules, we should be able to demonstrate at least three levels of difficulty merely by changing the way in which these eight stimuli are classified. Moreover, whereas the classification shown in Box I is a kind that has often been used to study the acquisition of concepts, the classification illustrated in Box VI might approach in difficulty a rote identification task in which a different response must be associated with each of the eight stimuli. (It might be noted that, if the rule given for this classification in terms of the logical connectives of conjunction and disjunction is expanded, it is seen to be logically equivalent to a complete enumeration of the four stimuli to which Response A has been assigned.)

Six Basic Types of Classifications

Fortunately not all 70 of the possible classifications need to be examined separately; for they belong to only six basic types. And any classifications belonging to the same type are essentially equivalent. For example, a classification that depends upon the value of only one dimension can be regarded as the same general type of classification whether the critical dimension is that of color (as in Box I) or that of size or shape. Likewise the decision as to which of the two classes

[2] The investigation of classifications of this kind (i.e., in which stimuli taking on one of two values on each of a small number of dimensions are divided into two equal classes) was probably in the *Zeitgeist*, judging by the number of independent (and often unpublished) studies of such classifications that were brought to our attention after completing the experiments reported here. Our own decision to undertake an exhaustive exploration of all possible classifications of this kind grew, in part, out of discussions with Alex Bavelas who was at that time a member of the technical staff of the Bell Telephone Laboratories. We are much indebted to him for interesting suggestions along these lines.

shall be assigned Response A and which Response B seems insignificant. Generally, then, we are led to say that two different classifications are of the same *type* if and only if one can be obtained from the other simply by interchanging the roles of the three dimensions or by reversing the two responses. In this way the 70 different classifications can be reduced to just six structurally distinct types. Each of the classifications shown in Figure 1 is an example of a different one of these six types. Accordingly, we shall henceforth refer to the corresponding types by the roman numerals I–VI. A detailed demonstration that there are just these six types will not be given here; it can be found in works on Boolean algebra and the theory of switching circuits (e.g., Higonnet & Grea, 1958, pp. 188–194).

In the experiments to be described many different kinds of stimuli were used, but (with one exception) all can be characterized in terms of three dimensions with two possible values on each. Hence we need a way of abstractly representing the eight stimuli and their six types of classifications without regard for the particular way in which the dimensions and values of the stimuli are realized, physically, in any particular experiment. A useful way of doing this is to set up a correspondence between the eight stimuli and the eight corners of a cube so

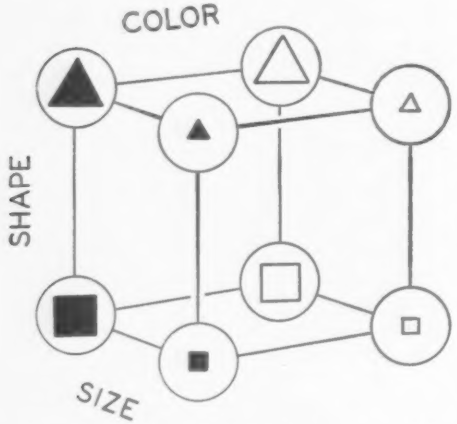

FIG. 2. An abstract representation of the eight stimuli as the eight corners of a cube.

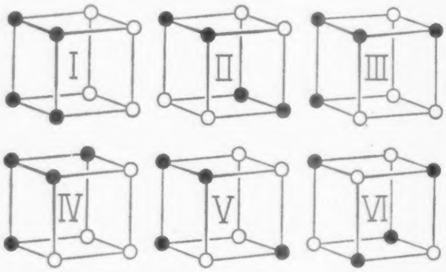

FIG. 3. The six basic types of classification represented abstractly by coloring four corners of the cube black and the remaining four white.

that the three dimensions of the cube represent the three dimensions of the stimuli. Such a correspondence is illustrated, for the stimuli of Figure 1, in Figure 2. As can be seen, the four stimuli having any one property in common (i.e., having the same value on one dimension) all fall on one face of the cube. Furthermore, stimuli having two properties in common are separated by a single edge, stimuli having one property in common are separated by two edges (or a face diagonal), and stimuli having no properties in common are separated by three edges (or a body diagonal).

Any of the 70 possible classifications of the stimuli into two equal subclasses can then be indicated by coloring four of the eight corners black and the remaining four corners white. The abstract representations for the six classifications illustrated in Figure 1 are shown in this way in Figure 3. Any of the other classifications can be obtained from one of these six simply by rotations and reflections of the cube. On the other hand, no two of these six can be obtained from each other by any combinations of rotations and reflections.

The types of classifications differ in that, in order to classify the stimuli correctly, they require knowledge of the values on only one dimension, for Type I; two dimensions, for Type II; or all three dimensions, for Types III–VI. Although these last four types are alike in that all three dimensions are relevant, they differ structurally in certain ways to be considered later.

The three ensuing sections describe in detail three experiments designed, first of all,

to compare the six basic types of classifications with respect to how difficult each type is to learn or remember. In Experiment I the stimuli are presented successively according to the usual paired-associate procedure (except, of course, that there are only two responses). The measure of difficulty in this experiment is the number of errors made during learning. Experiment I also provides information about transfer of classification learning since *S*s learn, in succession, several classifications of the same basic type but using different stimuli. In both Experiments II and III, on the other hand, the stimuli are presented simultaneously, as already grouped into the two classes. Then, following a period of inspection, *S*s either attempt to formulate a concise rule for how the stimuli can be sorted into the two classes or else attempt actually to so sort the stimuli. These two experiments also furnish information about how the way in which the dimensions and values of the stimuli are represented by the physical features of the stimuli affects the difficulties of the different types of classifications. The main variation there is between "compact" stimuli (like those in Figure 1) in which all three dimensions are represented by different aspects of the same object and "distributed" stimuli in which each dimension is represented by variations in a different one of three spatially separated objects.

After the summary of the empirical results of these three experiments (which immediately follows the detailed presentation of Experiment III) we attempt to evaluate alternative theoretical notions about classification learning with respect to their ability to account for the experimental results. In particular, we shall argue that neither the models of stimulus generalization nor those of the conditioning of cues are alone sufficient but that, in addition, something like abstraction and the formulation of rules is apparently involved.

Experiment I

The experiment to be reported first was designed primarily to answer two questions: How does the difficulty of learning vary from one type of classification to another? Is something specific learned about the structure of a classification that will transfer positively to the subsequent learning of a new classification of that same type? The experimental procedure conformed to the usual paired-associate paradigm except that only two responses were used. That is, the eight stimuli were presented, one at a time, in a continuing random sequence and an association between each stimulus and one of two alternative classificatory responses was built up by the method of anticipation. In order to obtain further information about the relation between identification and classification learning with the same set of stimuli, though, a condition was also included in which a different response was assigned to each of the eight stimuli.

Method

Subjects. Six female freshmen at Fairleigh Dickinson University served for 15 hours each in this first experiment. These *S*s were selected to be as uniform as possible with respect to their college entrance examination scores.

Learning tasks. Each *S* went through a different sequence of 27 learning tasks called *problems*. During any one problem *S* learned to associate a prescribed verbal response to each of eight stimuli. In most of the problems one response (e.g., A) was assigned to four of the eight stimuli and another response (e.g., B) was assigned to the remaining four stimuli. Except in certain special problems, each of the eight stimuli took on one of two values on each of three dimensions. Thus any type of classification from I through VI could be established. The stimuli were photographed on individual frames of 16-mm. film and projected onto a screen in front of *S*. A self-paced method of anticipation was used: i.e., as soon as *S* responded to a given stimulus she was told what the correct response for that stimulus in fact was and then the film was advanced to the next frame. Each of the eight stimuli on a given film occurred 50 times making a total of 400 frames. The order of the stimuli on each film was random except for the following constraints: within the first two blocks of 8 frames, each stimulus appeared exactly once; within every succeeding block of 16 frames, each stimulus appeared exactly twice. For each problem, learning continued until *S* attained a criterion of 32 consecutive correct responses.

Stimuli. Each of six film strips was prepared from a different set of eight stimuli. Figure 4 shows the eight stimuli used in one of these film strips. Each of the three positions in a stimulus represented a dimension in which either of two

thematically related drawings could appear as values. The arrangement of the eight stimuli in Figure 4 is analogous to the arrangement of the stimuli in Box I of Figure 1. Thus a Type I classification could be established by assigning Response A to the four stimuli on the left and Response B to the four stimuli on the right. In this case *S* would merely have to note whether the stimulus contained a candle or a light bulb in the lower left position in order to master the classification. These pictorial stimuli (rather than the simple geometrical figures of Figure 1) were selected for this first experiment in order to simplify the task of constructing the many different but comparable sets of stimuli required by the experimental design. Thus it was relatively easy to prepare another film in which either a desk or a chair appeared in the top position, an eye or an ear in the lower left, and a spool of thread or a pair of scissors in the lower right, and so on for the other films. One of the six films was constructed in a different manner. The same kinds of drawings were used, but an entirely different set of three

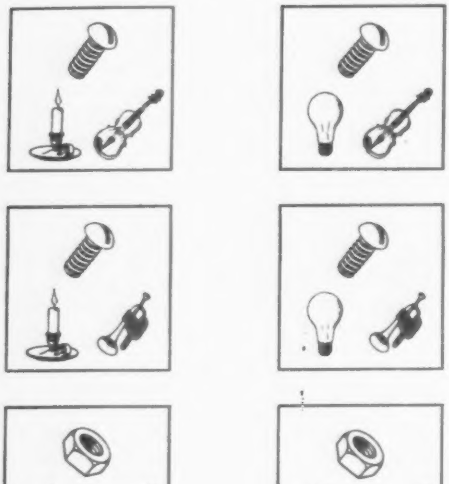

FIG. 4. One of the sets of eight stimuli used in Experiment I.

drawings was selected for each of the eight stimuli. Since no two of the eight stimuli had any picture in common, this is referred to as the "nonoverlap" film.

Experimental design. Table 1 presents the overall design. There were four general classes of problems. Two of these used the five films with "overlapping" stimuli (i.e., stimuli with two values on three dimensions); they were identification problems (ID_o, for "identification with overlap") in which a different response-letter (D, H, K, M, O, R, S, or W) was associated with each of the eight stimuli, and classification problems (I, II, III, IV, V, or VI) in which one of two responses was associated with each of four of the eight stimuli. Each classification problem used one of five alternative sets of responses: "A" and "B," "plus" and "minus," "P" and "Q," "one" and "two," or "X" and "Y." The other two general classes of problems used the nonoverlap film; they were identification problems ($\mathrm{ID}_{\bar{o}}$) in which the eight letters of the alphabet were assigned to the eight stimuli, and classification problems ($C_{\bar{o}}$) in which each of two responses was associated with four of the eight stimuli. (Since none of the stimuli on the nonoverlap film had any common properties, all $C_{\bar{o}}$ problems are of the same type.)

As indicated in the table, each of the six *S*s was given five consecutive problems of one type, then five consecutive problems of another type, and so on for four different types. A different film (i.e., set of stimuli) was used for each of the five classification problems of the first type administered to each *S*. The order of these films, however, was different for different *S*s. On subsequent types of problems the five films were used again for each *S* in the same order in which they had been presented during the five problems of the first type for that *S*. The order of the films for the identification problems (ID_o) was the same except that it began with the assignment of the first film to Problem 24 and ended with the assignment of the last film to Problem 1. The one nonoverlap film was used for both Problems 22 and 23. For each *S* the same set of responses was retained throughout the problems of a single type, but changed when the *S* proceeded to the next type. Indeed all *S*s had A and B as responses for the first five classification problems, plus and minus as responses for the next five, and so on. For each problem the assignment of the response to the stimuli was chosen at random from all possible assignments that would conform to the type prescribed by the design. Thus for a given *S* on the first problem of Type II the lower right drawing might be the irrelevant one whereas on the second problem of Type II the upper drawing might be the irrelevant one.

Every *S* learned five problems of each of the Types I, II, and VI but (owing to limitations of time) five problems of only one of the Types III, IV, and V. The sequence of types was different for each *S* but counterbalanced to the extent that each type occurred as frequently toward the begin-

ning as toward the end of the series of classification problems. (Unfortunately a complete counterbalancing could not be achieved with the number of Ss available. Hence there is a partial confounding of types and order of presentation of types.) Since the Ss could usually be scheduled for only an hour at a time, the following rules were adopted: If an S reached criterion on a given problem early in the hour, she was started on the next problem. If she reached criterion late in the hour, she was not given a new problem until the next session. And if (as occasionally happened) she had not reached criterion by the end of the hour, she was continued on that same problem at the beginning of the next session. All Ss appeared three times per week at roughly regular intervals throughout each week.

Instructions. The nature of the learning tasks was explained to each S, but nothing was said about the structures of the different types of classifications (I–VI). The experimenter (E) simply stated that S would learn five problems that were similar and of about the same level of difficulty, then five more problems that again were similar among themselves but different from the first five problems, etc. The statement was also made that the problems in each block of five might be more difficult or less difficult than the problems in the preceding block of five. Each S was urged not to discuss the problems with other Ss until after the experiment was completed. (Also, the variation of the order of the films, order of types of problems, and assignment of responses from one S to another presumably minimized the opportunity for communication of this kind.)

At the beginning of the first problem for each S the manner of construction of the eight stimuli was described. And, at the outset of that and each subsequent problem, the eight stimuli were shown one at a time. Before each problem S was also told what the set of responses for that problem would be. During identification learning S was not required to guess the responses for the first eight stimuli, but was simply told what these were. During classification learning, though, there were only two responses and S was required to guess from the outset. Each time a new problem was begun that was of the same type as the preceding problem, this fact was pointed out to S. Likewise, when the problem was of a new type, S was told that, although the stimuli would be the same as those she had already seen in an earlier problem, the problem itself might be quite different from the preceding problems.

After criterion was reached on any problem, E informally asked S whether she had any observations to report concerning the problem or how she had gone about learning the responses. In order to minimize the influence of this inquiry upon the subsequent behavior of the S in other problems, suggestions that S should be able to verbalize a rule relating the responses to the stimuli were avoided. For the same reason, if S had nothing or only vague observations to report, no attempt was made to press for further explanation. Consequently the record as to the rules formulated by Ss during learning does not provide detailed information about all subject-problem combinations.

Results

This section presents the results of Experiment I in detail. A general summary of the results of this experiment (as well as of the other two experiments, II and III) will be found in the section Discussion of Empirical Results which immediately follows the detailed presentation of the results of Experiment III.

Relations between measures of problem difficulty. The following four measures of performance were taken: the total time (in

TABLE 1

EXPERIMENTAL DESIGN

Ss	Successive problems							
	1	2-6	7-11	12-16	17-21	22	23	24-27
S_1	ID_0	I	II	(V)	VI	C_5	ID_5	ID_0
S_2	ID_0	VI	(III)	II	I	C_5	ID_5	ID_0
S_3	ID_0	I	VI	II	(IV)	C_5	ID_5	ID_0
S_4	ID_0	(IV)	II	VI	I	C_5	ID_5	ID_0
S_5	ID_0	II	(V)	I	VI	C_5	ID_5	ID_0
S_6	ID_0	VI	I	(III)	II	C_5	ID_5	ID_0

Note.—Problems 2-21 constitute the main sequence of classification problems. During this sequence each S was given each of the three Types I, II, and VI, but only one of the three Types III, IV, and V. These latter types are enclosed in parentheses in order to set them apart in the table.

minutes) required to reach criterion (t), the total number of stimulus presentations excluding the 32 presentations within the criterion sequence itself (p), the total number of incorrect responses (errors) made prior to reaching criterion (e), and the number of errors during the first 32 presentations (f). The correlations between these measures (over-all 162 subject-problem combinations) were as follows: $r_{tp} = 0.93$, $r_{te} = 0.94$, $r_{tf} = 0.67$, $r_{pe} = 0.90$, $r_{pf} = 0.63$, and $r_{ef} = 0.77$. The f measure was probably the least reliable of the four since it was based on only the first 32 presentations. This may account for its low average correlation with the three other measures (viz., 0.69). Nevertheless, all four measures were found to have essentially the same relations to the independent variables of the experiment. Rather than carry along all four, then, only the total number of errors, e, will be used as the dependent variable in what follows. This measure had the highest average correlation with the other three (viz., 0.87) and has been used in previous studies of classification learning (cf. Bourne & Restle, 1959; Smith, 1954).

Transfer of classification learning. Figure 5 presents the mean number of errors per S on the 20 principal classification problems (Problems 2–21 of Table 1). Beyond the over-all (secular) decline throughout the 20 problems, there was a pronounced drop in

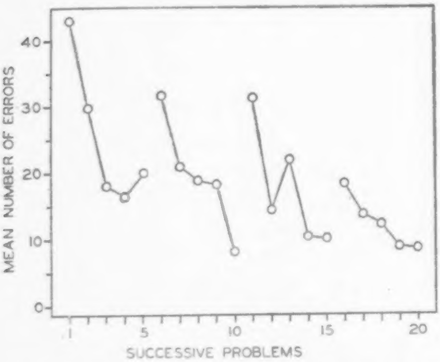

Fig. 5. Mean number of errors for the main sequence of 20 classification problems. (Problems 2–21 in Table 1. Points corresponding to problems of the same type are connected by lines.)

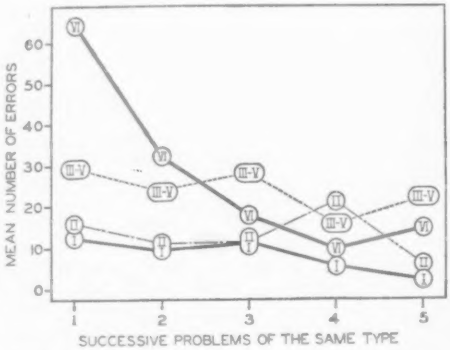

Fig. 6. Mean number of errors for the first through fifth classification problems of each type. (Separate curves are plotted for Types I, II, and VI, but a single curve for Types III, IV, and V combined.)

errors during each block of five consecutive problems of the same type. This specifically within-type transfer occurred even though the stimuli and the assignment of responses changed for each new problem. What transferred, then, was neither a particular set of stimulus-response associations nor simply a generalized increase in ability to handle new problems; it was something about the unique structure of the type itself.[3] However, as will be seen, most of this within-type transfer can be attributed to Type VI classifications alone.

Comparisons between the six types of classifications. Figure 6 shows how the over-all difficulty and the within-type transfer varied from one type of classification to another. The individual curves for Types III, IV, and V are not presented separately.

[3] Also, of course, the sudden increase in errors whenever a new type of classification was introduced might in part be a consequence either of an emotionally disrupting effect of the instruction that the next problem would be of a new kind, or else of some interference resulting from the repetition of the set of stimuli used five problems earlier with different responses. However, these explanations seem implausible in view of the presence of a strong interaction between positive transfer and type of classification, the total absence of overt intrusions of previously correct responses during the new problem, and the subsequent reports of the Ss themselves.

Since they were based on only two *S*s each, they are quite erratic and their inclusion would therefore tend to obscure any pattern in the more stable curves. For this reason and because theoretical considerations indicated that these three types might be about equal in difficulty, they were averaged together to yield a single curve that is comparable in stability with the curves for Types I, II, and VI.[4]

The reliability of the pattern exhibited in Figure 6 was evaluated as follows: First, the number of errors, e, made by each S on each of the 20 classification problems was transformed to yield a new difficulty score e' by the commonly used logarithmic transformation $e' = \log_2 (e + 1)$. The log transformation (used also in Smith's study, 1954) largely eliminated the initially apparent dependence of the variance of errors upon their mean value. Then, a second-order orthogonal polynomial was fitted to the five transformed points in each one of the 6×4 cells corresponding to a different subject-type combination. Thus the set of five error scores in a given cell was reduced to a set of three coefficients: the mean value of e' for the given S on the five problems of the given type, the linear trend in e' over the five problems, and the quadratic curvature of that trend. Finally, an over-all analysis of variance was carried out on these coefficients.[5]

The analysis indicated that the four curves in Figure 6 do differ reliably in over-all level ($F = 4.5$, $p < .05$), in linear trend ($F = 8.0$, $p < .01$), and in curvature ($F = 6.4$, $p < .01$). The average left-to-right trend of all these curves taken together is also statistically significant both in its linear component ($F = 26.4$, $p < .01$) and in its curvature component ($F = 8.4$, $p < .05$). However, as is clear from Figure 6, this average linear trend and curvature as well as the differences between types with respect to trend and curvature is largely attributable to the very marked decline and concavity of the curve for Type VI alone. Order of presentation of types of classifications had no significant effect upon the height of the curve for each type ($F = 1.0$), the linear trend ($F = 3.4$), or the curvature ($F = 3.0$). The fact that the left-to-right linear trend in Figure 6 is significant whereas the effect of order of presentation of types is not significant further supports the conclusion, drawn from Figure 5, that the within-type positive transfer was reliably greater than the between-type positive transfer. In a somewhat more rigorous test of this point, the over-all downward trend of the curves in Figure 6 was found to be significant even after the linear component of the order effect (i.e., the secular decline in Figure 5) was subtracted out. (Again, however, this within-type trend is largely contributed by Type VI alone.) Finally, the six *S*s did not differ significantly either in over-all performance ($F = 2.4$) or in the linear trend of their performance over a series of problems of the same type ($F = 1.7$). Oddly, however, they did differ in the quadratic component of this trend ($F = 8.3$, $p < .01$). This last result appears to be primarily attributable to one S (viz., S_1) whose curves were all convex (rather than concave) upwards.

Since the four curves presented in Figure 6 evidently do differ reliably, a more detailed examination of these differences was

[4] Unfortunately the experimental design does not permit an adequate test of possible differences between the individual curves for Types III, IV, and V. With the data that were obtained, however, the curves for III and V did not exhibit any consistent differences either in over-all height or in trend. The curve for IV did generally fall somewhat below the curves for III and V; but the two *S*s who were given Type IV problems happened to be the two who made the smallest number of errors on the other problems also. About all that can be claimed at this point is that the results are at least consistent with the conclusion (established more firmly in the following experiments) that Types III, IV, and V are essentially equal in difficulty.

[5] The fact that the experimental design entailed a partial confounding of types and order of presentation of types necessitated the computation of certain correction coefficients in order to render orthogonal any comparisons involving these two variables. We are greatly indebted to M. J. R. Healy who, while a visiting member of the Bell Telephone Laboratories, proposed the method of analysis and derived the formulas for the required correction coefficients.

undertaken. With respect to the first problem learned of each type (the left-most point of each curve), the results of the comparisons between types (using two-tailed t tests) were as follows: the first point for VI was significantly different from the point for III, IV, and V combined ($p < .01$); the point for III, IV, and V, in turn, was different from the point for II ($p < .05$); the difference between the first points for II and I, however, was not significant. The differences in the initial difficulties of the six types cannot be attributed to interference or negative transfer from preceding classification problems; for the same ordering is found when we look at the first problem of only the first type learned by each *S*. Thus the average number of errors made on the very first classification problem (Problem 2 in Table 1) was 7, for Type I; 29, for Type II; 41, for Type IV; and 86, for Type VI. (None of the *S*s had Types III or V first.)

The comparison of the differences between types for the points corresponding to the second through fifth problems of each type is complicated by the significant dependence of within-type trend upon type. However, for the purposes of the subsequent theoretical discussion, it is sufficient to observe that the downward trend for VI is significantly greater than that for any other curve. Thus the results are consistent with the hypothesis of an initial ranking: I < II < (III, IV, V) < VI. But they also indicate that, with continued practice, VI decreases in difficulty relative to the other types and, so, eventually becomes easier than III, IV, and V (considered together).

Analysis of rules verbalized by Ss. For all but four of the subject-problem combinations *S* described the classification in terms of an explicit rule. All but five of these explicit rules proved to be correct in the sense that, by sorting the stimuli in accordance with the given rule, *E* could reconstruct the correct classification. All of these rules were rated for amount of unnecessary complexity; i.e., the amount of complexity of the stated rule over and above that of the least complex rule possible for the given type of classification. Some of the types admit two different rules that seemed about equally economical. Examples of each of the kinds of rules that were taken to be most economical for each type of classification are as follows:

Type I: "If there's a candle it's an A; otherwise B."

Type II: "If there's a candle and trumpet or if there's a light bulb and violin it's an A; otherwise B."

Type IIIa: "If there's a candle but not both a violin and screw, or if there's a violin and nut, it's an A; otherwise B."

b: "If there's a candle and trumpet or if there's a violin and screw it's an A; otherwise B."

Type IVa: "If there's a candle but not both a violin and screw, or if there's a trumpet and nut, it's an A; otherwise B."

b: "If there's a candle, violin, *and* screw or any two of these it's an A; otherwise B."

Type V: "If there's a candle but not both a violin and screw, or if there's a violin and screw but not a candle, it's an A; otherwise B."

Type VIa: "If there's a candle, violin, *and* screw or just one of these three it's an A; otherwise B."

b: "If either just one or else all three pictures change the response changes to the other alternative; otherwise the response remains the same."

The rule for Type I simply specifies the values on the one relevant dimension. The rule for II does the same for the two relevant dimensions. The rules for IIIa, IVa, and Va are of the same general kind. They might be called "single dimension with exceptions" rules in that they specify not only the values on one relevant dimension (as in Type I) but also the two exceptional stimuli for which the responses must be reversed.[6] The rule for IIIb is similar to that for IIb except that three rather than two dimensions

[6] In general this kind of rule is of this form: "If there's a candle it's an A; otherwise B; except the one made up of a candle, violin, and screw must be exchanged with the one composed of a light bulb, trumpet, and nut." The specific forms of this "single dimension with exceptions" type of rule that were given above differ from this general form in that they take advantage of certain subtle differences in the structures of Types III, IV, and V in order to shorten the rules slightly by omitting mention of one value for each of the two exceptional stimuli. However, *S*s were considered to have formulated the simplest rule for Types III, IV, or V even if they mentioned all three values of the exceptional stimuli as in this general form of the rule.

are involved. The vocabulary is coordinate for all the rules just considered; but for the further rules IVb, VIa, and VIb—which we shall sometimes refer to as the "odd-even" rules—a new process of counting enters. (As indicated in the introduction, by the length of the rule for VI given there in terms of a logical conjunction and disjunction of values, the rule for this type is extremely complicated if counting is not used.) The rules IIIb and V[illegible] are similar in that the classification is defined in terms of the three pictures contained in a single "pivotal" stimulus. However, they differ in that, whereas any stimulus can be chosen as pivotal in VI, only two of the eight stimuli will serve in III. The final rule, VIb, is unique in that it involves a comparison of two consecutively presented stimuli and, so, would not provide any basis for responding to the first stimulus presented.

Three judges (JG, HMJ, RNS) independently rated the amount of complexity of the rules stated by the *S*s over and above that of these most economical rules. The rating was done in random order with knowledge of the type of classification involved but without knowledge about which of the six *S*s produced the rule or how many problems had already been learned by that *S*. The rating was done on a five-point scale according to the following guide:

0. Equivalent to one of the most economical statements of the given type of classification.

1. As above, but either includes the complete rule for both responses (rather than simply specifying the application of the second response by exclusion) or else states the values on an irrelevant dimension, but not both of these. (Example of stating an irrelevant dimension in Type I: "If there's a candle and a violin or trumpet it's an A; otherwise B.")

2. Fails to use exclusion and includes irrelevant dimension; or repeats some information unnecessarily; or is incomplete.

3 and 4. Increasing degrees of complexity.

5. Enumerates all four stimuli in each class or states a rule judged to be equivalent to such an enumeration in terms of the number of dimensions and values specified.

The three pair-wise correlations between the three judges' ratings of the amount of unnecessary complexity of the stated rules were .94, .90, and .89. Since the judges seemed to produce similar ratings, their ratings were averaged for comparisons with other variables. Their mean ratings of unnecessary complexity for Types I, II, III, IV, V, and VI were 1.6, 1.7, 2.9, 1.8, 3.2, and 2.3 (in that order). For purposes of comparison, the number of errors made during the learning of classifications of each type (averaged over all five problems of the same type) were 8.3, 13.2, 32.0, 16.7, 23.6, and 28.0, respectively. The correlation between these two sets of numbers is statistically significant ($r = .80$, $p < .05$). There was also a decrease in the rated complexity of rules in successive problems of the same type. This is shown, along with the corresponding reduction in errors (for all types taken together), in Figure 7. The similarity of the two curves suggests that there may be a close relation between the reduction of errors and the discovery of a more economical rule. Type VI showed the greatest reduction in complexity of stated rule, just as it showed the greatest reduction in errors, during the course of the five successive problems.

Identification and nonoverlap conditions. In all conditions there were only eight different stimuli; and these were highly discriminable in the sense that the difference be-

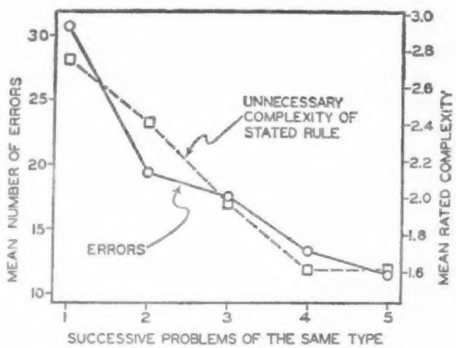

FIG. 7. Mean number of errors and mean rating of the amount of unnecessary complexity in the rule stated by *S*s for the first through fifth problems of the same type. (The average is taken over all six types of classifications. The location of the zero points and the size of the units have been adjusted to bring the two curves into approximate alignment.)

tween any two could easily be seen and described by *S*s. Moreover these stimuli were meaningful in the sense that they depicted familiar objects for which the *S*s already had overlearned verbal labels. Now a paired-associate list composed of only eight highly discriminable and meaningful stimuli would seem to be relatively easy. Yet the six *S*s found most of the classification and identification problems quite difficult. On the average, they mastered the first identification problem (ID_o) only after 59 minutes, 249 stimulus presentations (which is equivalent to 31 times through the "list"), and 103 incorrect responses. The reason for the difficulty of this task is made clear by a comparison with the corresponding nonoverlap condition ($ID_{\bar{o}}$). This problem was mastered, on the average, after only 3.7 minutes, 4.7 stimulus presentations, and 0.8 incorrect responses. Of course, of these two identification problems, the one with overlapping stimuli always preceded the one with nonoverlapping stimuli, but this cannot alone account for the hundredfold difference in errors. For the average number of errors on the last four identification problems with overlapping stimuli (ID_o) was still 23 (as opposed to 0.8 for the nonoverlapping stimuli).

The results, then, support the following account: The identification problem with nonoverlapping stimuli was relatively easy because *S* could associate the response for each stimulus with each of the component pictures of that stimulus independently. In fact, the problem could be mastered by attending to the picture in only one (say the lower left) position. And the pictures appearing in this position were indeed highly meaningful. Owing to the overlap in the component pictures in the contrasting identification problem, however, *S* could never achieve a satisfactory performance by attending to the picture in a single position. Rather, *S* would have to learn to identify the unique pattern corresponding to each of the eight combinations of three pictures. And, although the individual pictures were meaningful, their combinations were not. Moreover, the overlearned verbal labels (e.g., "candle," "violin," etc.) were of little help; stimuli that shared pictures would also share these labels. A verbal response uniquely associated with each combination of three pictures presumably would have been helpful, but this *S* did not initially have. The situation here is analogous to a rote learning task with nonsense syllables as stimuli. There, *S* has an overlearned response to each individual letter, but not to the whole pattern of three.

The comparison between the identification and classification problems with nonoverlapping stimuli is unfortunately confounded completely with order effects; for, whereas the same set of nonoverlapping stimuli was used for both conditions, the classification problem always preceded the identification problem. Still, in view of the absence of a strong order effect among the other classification problems, it seems rather surprising that the classification problem with nonoverlapping stimuli resulted in an average of 6.7 errors as opposed to the average of only 0.8 errors for the corresponding identification problem.

Experiment II

This experiment was designed to secure more systematic information concerning the kinds of rules spontaneously formulated by *S*s in categorizing stimuli arranged according to the various types. In addition, information was sought on the following problems: Is the difficulty of memorizing the classification into which the stimuli have been arranged related to the difficulty of formulating rules for their classification? Does the way in which the dimensions and their values are represented by features of the stimuli affect the difficulty of memorization or formulation of rules for a classification?

These problems were investigated through the use of two tasks: *rule formulation,* in which *S*s were first exposed to eight stimuli divided into two groups of four each on the basis of one of the various types of classifications discussed above and then asked to give the basis on which the two sets could be differentiated; and *memorization*, in which the same *S*s were subsequently (1–2 weeks later) exposed to the same divisions of the

stimuli into two groups but with instructions to study them until they had memorized the classification, and then were tested for their ability to sort the stimuli correctly into the two groups.

Method

Subjects. Data were secured from 20 college students from an elementary psychology class at the University of Bridgeport. Four other students were dropped because of failure to follow the instructions on preliminary trials.

Stimuli. Classifications of Types I, II, III, V, and VI were employed. Three different sets of each type, except VI, were chosen in such a way as to counterbalance for each type, the roles played by the three dimensions. (Thus, for each of the three Type I problems, a different dimension would be the relevant one.) Only one problem of Type VI was given since all three dimensions play the same role in this particular type. Altogether $3 \times 4 + 1$- or 13 different sets of eight stimuli were therefore utilized.

The three dimensions and two values of the stimuli were represented by stimulus features in three different ways. These are illustrated in Figure 8 where (for each of the three ways) the stimuli are arranged according to a Type II classification (with the four stimuli on the left in one class and the four on the right in the other class).

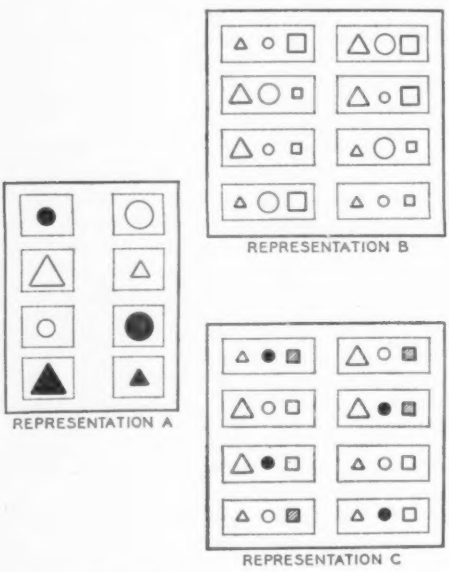

FIG. 8. Three different perceptual representations of a Type II classification.

We refer to the first way (A) as a "compact" representation since each stimulus is a single geometrical figure. The three dimensions of the stimulus are size, color, and shape. The corresponding values are large-small, black-white, circular-triangular. The second and third representations (B and C) are referred to as "distributed" since each stimulus consists of three separate geometrical figures (triangle, circle, square) each of which is used to encode one of the three dimensions. In B, each figure may be large or small giving us the two values for each dimension. In C, however, the values are represented in a different way for each figure. Thus, the triangle may be large or small, the circle may be black or white, and the square may be shaded or unshaded. (This last kind of representation, C, is presumably the most closely analogous to the stimuli used in Experiment I, which were all "distributed" in the present sense.)

The choice of these three perceptual representations was not based upon any well-defined theoretical predictions but was intended to provide some generality of materials and also to evaluate certain a priori expectations as to how the type of coding of the dimensions might affect performance. It was thought, for example, that the compact Representation A might differ from the distributed representations in that a single discrete response might be made to the single geometrical figure without having to respond successively to each of three spatially separated figures. Illustrative of another possible influence of the perceptual representation would be the expectation that the Type VI problem would be easiest with Representation B, since here this otherwise difficult problem could be reduced to an odd-even rule based simply on the number of large figures. A final example is the possible formulation of a rule for Type II in terms of the "same" or "different" size of two of the figures (e.g., "when triangles and squares are same size, put on the right," etc. in Figure 8). It will be found later, however, that the differences between the two distributed representations were not generally of significance, and in most of the subsequent discussion the results for these two representations (B and C) will be combined. The nature of the dimensions and values for the three different representations are specified in Table 2.

Formulation of rules. The first task for each *S* was that of formulating rules for categorizing the stimuli presented to him. Instructions were designed to exclude simple enumeration in terms of a complete specification of all values for each of the eight stimuli.

The test booklet containing an arrangement of the eight stimuli into one of the 13 classifications was presented to *S*. The *E* had a set of eight cards, one for each stimulus, that he was prepared to sort into two piles on the basis of the rule given to him by *S*. The *S* was instructed as follows:

I have a duplicate set of the cards which have been presented to you in two groups in booklet form. What I want you to do is to tell me which

TABLE 2

NATURE OF DIMENSIONS AND VALUES FOR THE THREE DIFFERENT PERCEPTUAL REPRESENTATIONS

Perceptual representation	Dimension 1	Values		Dimension 2	Values		Dimension 3	Values	
		1	2		1	2		1	2
A	Color of figure	white	black	Size of figure	small	large	Form of figure	triangle	circle
B	Size of triangle	small	large	Size of circle	small	large	Size of square	small	large
C	Size of triangle	small	large	Color of circle	black	white	Shading of square	open	shaded

figures belong in the A group, and which figures belong in the B group. You are not allowed to describe particular cards. You must give general descriptions of categories or classes of cards. [*E* gives an example.] This is not an intelligence test; we are interested in how individuals describe sets of items.

Do you have any questions at this time regarding what you are to do? If not, we'll start with the first set of items.

The exact rules given to *E* by *S* for categorizing the cards were recorded. A sample protocol for Type I might read: "Put all the large figures on the left and all the small on the right." The time elapsing between the opening of the test booklet and the statement of a rule by *S* was recorded. A second *E* recorded *S*s rules and the time required for their formulation.

If *S* gave a rule by which *E* could sort his set of cards, the rule was followed and recorded. If the rule was not clear or violated the instructions given to *S*, *E* challenged *S*, explaining that all the stimuli were not covered or that *S* was not following instructions—particularly when *S* only described the individual cards. The nature of the challenge and *S*'s response to it were recorded.

Memorization. Two weeks later each *S* was given a test of his speed and accuracy in memorizing the assignment of stimuli according to the various types of classifications. In this test *S* was given a set of eight cards constituting the total set of stimuli. *E* then presented to *S* one of the 13 booklets described above. *S* was instructed to study the booklet until he felt he could distribute his set of eight cards into two categories (left and right) in the same way as they were distributed in the booklet. His performance was timed and the distribution of the cards into the two piles was recorded. Thereupon another set of instances was presented to *S* by *E*. The instructions were as follows:

For this phase of the experiment you will be shown a booklet which contains eight cards, divided into two groups. When you feel that you are ready to sort your own cards tell me. I'll close the booklet and you pick up your cards and sort them into the same two groups. The order in which the cards appear (within each group) is not important. Just get the same cards into the same grouping.

Don't pick up your cards until I have closed the booklet. Do you understand the task? Any questions? All right, let's begin.

E opened one of the test booklets and, when *S* indicated he was ready, *E* shut the booklet and *S* sorted the eight cards into two piles from memory. The sorting was scored only on the basis of getting the stimuli correctly grouped together; whether one pile was placed on the right or left was not taken into account. (Since the order of the two groups was not scored, there were only 35 rather than 70 possible classifications here.) The time elapsing between the presentation of the instances to *S* and the beginning of his sorting of the stimuli was recorded to the nearest second.

Experimental design. In both the rule formulation and the memorization tasks each *S* received all 13 different sets of instances in each of the three perceptual representations. The three representations were contained in a latin square with approximately one-third of the *S*s receiving the compact Representation A first, one-third Representation B, and one-third Representation C. *S*s for each of the three groups were assigned randomly. A similar latin square was employed for the memorization task except for the restriction that no *S* had the same order of representations in the two tasks. Within a particular kind of representation the 13 sets of problems were given in random order.

Prior to the beginning of the experiment proper a series of trial tasks was employed that had substantially different geometrical figures but involved essentially the same kind of classifications as those used in the main experiment.

Results

Formulation of rules. While a number of different kinds of rules could be formulated for these materials comparable to those discussed under Experiment I, only three kinds constituted the bulk of the formulations actually produced by *S*s. Although it was forbidden in the instructions, enumeration of stimuli was occasionally used—primarily in connection with Type VI. Examples of the various types of rules may be helpful. The following examples are taken from those given for the compact perceptual Representation A:

Single factor: "All circles on the left; triangles on the right."

Two-factor: "Small triangles and/or large circles on the left; large triangles and small circles on the right."

Single factor with exception: "All black figures on the left and white figures on the right *except* the large black circle should be exchanged with the small white triangle."

It will be recalled that the simplest rule for Type I is the single factor rule; the simplest for Type II is the two-factor rule; for Type III either a two-factor formulation (involving three dimensions) or a "single factor with exception" rule is possible; the simplest for Type V is the "single factor with exception" rule.

Results are shown in Table 3. The percentages of *S*s giving a correct rule of each kind for the various types of problems are presented.[7] The sets representing the same problem but with different dimensions utilized as the basis for classification are indicated by 1, 2, and 3. General rules which were correct but not included as one of the three most common kinds are tabulated under "miscellaneous," and correct statements of the classifications by complete descriptions of the four stimuli in each group are tabulated under the column "enumeration." "Errors" involve an incorrect statement of the rule(s) or a failure to formulate a rule.

[7] The assignment of the rules produced by *S*s to the various categories of Table 3 was based upon agreement between one of the *E*s (CIH) and Albert Bregman (whose assistance as a judge is gratefully acknowledged).

The principal phenomenon is the high relative frequency with which *S*s formulated appropriate simple rules (numbers enclosed in boxes). But the extent to which the simplest rule was produced varied greatly with type of classification and kind of perceptual representation. (The fact that it also varied somewhat from one set to another within a given type and perceptual representation may mean that the three dimensions differed somewhat in salience. However, this kind of variation is generally small compared with that attributable to type of classification.) For Type I from 75–100% of the *S*s (with an exception to be noted at the end of the paragraph) formulated a single-factor rule. The extent of use of the simplest rule was actually greatest (though not significantly so) for Representation C, where a different figure and kind of variation was used to represent each dimension. It will be noted from the table that some *S*s utilized a less efficient rule for Type I in which they described the stimuli in terms of *two* dimensions rather than the *one* required. For example, instead of specifying "All circles on the left and all triangles on the right," they would say: "All large circles and all small circles on the left, and large triangles and small triangles on the right." This accounted for 50% of the formulations in the second of the three Type I problems with the compact Representation A.

With the compact Representation A for Type II from 65–90% of *S*s stated a simple, appropriate rule, one involving two factors. Thus almost as many *S*s formulated a simple rule for Type II as did so for Type I (with Representation A). The corresponding percentages for Representations B and C were substantially lower (30–55% correct) than for A, and were less than the corresponding percentages for Type I. With these representations *S*s appear to have considerable difficulty in discovering a rule for Type II classifications.

Results for Types III and V were closely similar to each other. There was a substantially smaller number of *S*s using the simplest applicable rule than in the case of Types I and II. There were no clear differences among the three kinds of representations.

TABLE 3

TYPES OF CLASSIFICATION RULES USED FOR THE SIX TYPES OF PROBLEM
(Percentage of *S*s utilizing each type of rule)

Problem Type	Set	Single factor	Two factor	Single factor with exception	Miscel-laneous	Enumer-ation	Errors
		Perceptual Representation A (Compact)					
I	1	85	15			—	—
	2	40	50			5	5
	3	85	5			5	5
II	1		65		20	—	15
	2		90		—	10	—
	3		85		—	—	15
III	1		10	35	15	—	40
	2		25	15	5	—	55
	3		10	40	5	10	35
V	1			30	35	—	35
	2			10	15	15	60
	3			30	25	—	45
VI						30	70
		Perceptual Representation B (Distributed: same values)					
I	1	75	20			—	5
	2	95	—			—	5
	3	75	20			5	5
II	1		30		10	5	55
	2		45		—	5	50
	3		55		—	5	40
III	1		5	45	—		50
	2		5	20	5		70
	3		15	15	5		65
V	1			40	—		60
	2			40	—		60
	3			35	10		55
VI							100
		Perceptual Representation C (Distributed: different values)					
I	1	100	—				—
	2	90	5				5
	3	100	—				—
II	1		45		—	5	50
	2		35		—	10	55
	3		50		5	—	45
III	1		10	35	—		55
	2		10	45	—		40
	3		5	35	5		60
V	1			20	5		75
	2			20	—		80
	3			35	—		65
VI						10	90

The substantial number of "miscellaneous" classifications primarily comprises ones in which two of the four stimuli on each side were classified by a general rule and the remaining two were then individually described.

No *S* in any of the groups gave a correct general rule for Type VI.[8] In violation of the instructions some of the *S*s enumerated the specific stimuli which were assigned to each group. Although these enumerations were contrary to the instructions, they were accurate statements of the classification for 30% of the *S*s with Representation A, 10% with C, and 0% with Representation B.

These results are summarized in Figure 9. It will be seen that there was an over-all progression of reduced adequacy and accuracy of the formulated rules with increased complexity of the classification. Accuracy was greater for the more complex classifications when the dimensions were perceptually represented in compact form (A) than when the information was distributed over three figures (as in B and C).

The length of time required to decide on an appropriate formulation was also analyzed. These times show essentially parallel results, with longer formulation times for the more complex types of classifications. There is one interesting inversion, however, in that *S*s took longer to formulate a rule for Type II than for III or V when perceptual Representations B and C were used. With these representations the rule for Type II seems extremely difficult to formulate. The over-all phenomenon, however, is that *S*s do not merely take a longer time to decide on an appropriate formulation but, even after this longer time, they still are less accurate in formulating the appropriate rules.

Memorization. Results for the second task, that of remembering the category to which each of the eight stimulus cards had been assigned, are presented in Tables 4 and 5. These data show a close relationship between the type of classification and the accuracy (Table 4) and speed (Table 5) with which the assignment of stimuli is memorized. Practically all *S*s correctly sorted the stimulus cards into the appropriate groups after a 2–5 second period of inspection in the case of Type I classifications. Consistent with the previous results on adequacy of formulation of rules, the greatest accuracy is attained with perceptual Representation C. Performance with Type II problems was substantially poorer than for Type I and the difference was most marked for perceptual Representations B and C.[9] Measures in terms of accuracy and those in terms of time for memorization are closely parallel.

Problems of Type III and V did not differ among themselves but both were clearly more difficult than Type II. Again the differences between Type II and Types III or V were more marked with Representations B and C.

[8] The fact that none of the *S*s in Experiment II discovered the highly efficient odd-even rule that was discovered by some of the *S*s in Experiment I is probably a consequence of two differences between the two experiments. First, the much larger number of learning trials for each problem as well as the consecutive presentation of problems of the same type provided much greater opportunity for the discovery of more effective rules in the first experiment. Second, the successive presentation of stimuli used only in the first experiment may have made it more likely that *S*s would notice how many (e.g., whether an odd or even number of) values changed from one stimulus to the next.

[9] None of the available statistical procedures is completely appropriate for frequency data (i.e., frequencies of correct formulation of rules or frequencies of correct memorization) when these frequencies are obtained from the same *S*s under different conditions. The analysis mentioned below, however, revealed no transfer from one perceptual representation to another or from one type of problem to another. Under these conditions the type of chi square analog of analysis of variance developed by Lancaster and described by Sutcliffe (1957) provides conservative estimates of significance since the remaining factor, individual differences, would certainly result in a still smaller error term. The difference between perceptual representations, according to this analysis, is significant at the .001 level. (This difference is primarily attributable to the difference between the compact representation, A, and the distributed representations, B and C, taken together.) The difference attributable to type of classification is significant at the .001 level also. The internal consistency of the data as well as their similarity to those of Experiments I and III attest further to their reliability.

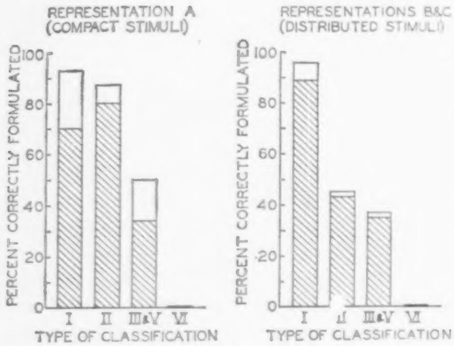

FIG. 9. Percentage of *Ss* formulating correct rules for each type of classification. (Shading indicates percentage formulating most efficient rule.)

Still more difficult in terms of accuracy and time for memorization is Type VI. The distributed perceptual representations (B and C), composed of three figures, once more resulted in greater difficulty than the representation in compact form (A). The difference in accuracy score between Representations B and C was in the direction predicted from the greater availability of the odd-even rule (discussed above) for Representation B, but the difference between one and four (out of a possible 20 *Ss*) is not statistically significant and we did not secure confirming evidence for its utilization in the rule formualtion task.

The results for accuracy and speed of memorization are shown graphically in Figure 10 where the data for the three kinds of perceptual representations are combined. The same relationship between accuracy and inspection times that was obtained in the case of the rule formulation task also was obtained here; i.e., accuracy was lower for the difficult problems even after additional time had been spent in assimilating the information.

TABLE 4

ACCURACY IN MEMORIZATION TASK: EXPERIMENT II
(Percentage of *Ss* correctly assigning stimuli from memory)

Problem Type	Set	Perceptual Representation A (Compact)	B (Distributed: same values)	C (Distributed: different values)
I	1	100%	100%	100%
	2	100	80	100
	3	90	85	100
II	1	85	60	70
	2	70	60	70
	3	90	75	80
III	1	65	50	35
	2	60	60	50
	3	65	30	50
V	1	60	35	45
	2	55	60	50
	3	85	30	35
VI		55	20	5

TABLE 5

TIME REQUIRED TO MEMORIZE STIMULI: EXPERIMENT II
(Median number of seconds elapsing between presentation of stimuli and *S*'s sorting of perceptual stimulus cards)

Problem Type	Set	Perceptual Representation A (Compact)	B (Distributed: same values)	C (Distributed: different values)
I	1	2	4	4
	2	4	5	4
	3	4	4½	3
II	1	7	21	23
	2	9	23	27
	3	7½	17	20
III	1	11½	21½	21
	2	12	24½	31
	3	11	23½	27½
V	1	12	27½	22½
	2	10	19	26
	3	9	27½	31
VI		18½	51	42

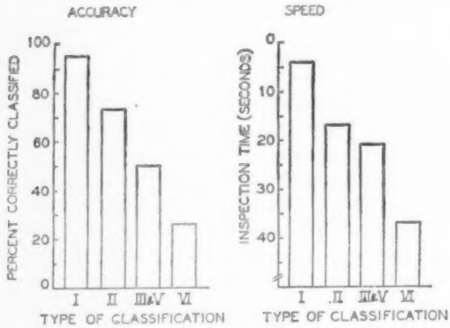

FIG. 10. Accuracy (percentage of *S*s correctly assigning stimuli to appropriate category) and speed (time required to memorize assignment) for various types of classifications.

The difference between the perceptual representation in compact form (A) and that distributed over three figures (B and C) is summarized graphically in Figure 11. The differentiation between types of classifications will be seen to be more marked in the case of the distributed Representations B and C than in the case of the compact Representation A.

Analysis was made of performance when the same task was carried out during the first, second, and third portions of the testing cycle. No significant improvement was found attributable to prior experience with the problems presented in other perceptual representations. Similarly, no improvement on successive problems within a particular kind of representation was found.

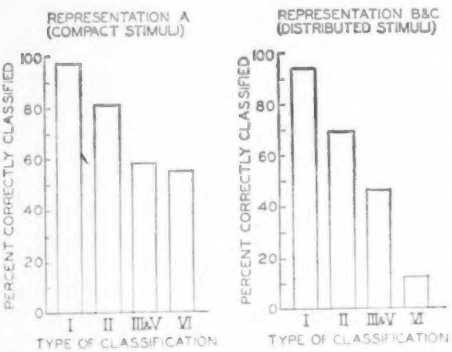

FIG. 11. Percentage of *S*s assigning stimuli to correct category for various types of classifications with compact (A) and distributed (B and C) perceptual representations.

Relationship between difficulty of formulation of rules and of memorization. The data presented in Table 3 were correlated with the corresponding results in Table 4 to determine the relationship between the extent to which the simplest rule was formulated in the first task and the accuracy of sorting from memory in the second task. A correlation was computed for 39 entries; i.e., for the 13 pairs of values for each of the three perceptual representations. The correlation was .90 and the regression was linear. This appears to indicate that a common factor is responsible for the increase in difficulty of rule formulation and of memorization.

EXPERIMENT III

It might be contended that the differentiation in memory scores for the various classifications of stimuli in Experiment II was attributable to the prior experience of formulating rules for classifying these stimuli. As a check on this possibility the memorization task alone was repeated with a new group of *S*s having no prior experience in formulating rules for classifications. Otherwise, essentially the same procedures were used.

Method

The procedures for the memorization task were identical to those described for Experiment II except that all six types of classifications (including IV) were employed. Two sets of Types I, II, III, IV, and V and one of Type VI were memorized by each *S*. The three perceptual representations illustrated in Figure 8 were again used.

The *S*s were 26 students from the elementary psychology course at the University of Bridgeport.

Results

Data concerning accuracy in memorization of the six types of classifications are given in Table 6. Corresponding results on time required to memorize the classifications are presented in Table 7. The increase in difficulty with increased complexity of type of classification is again evident. Type I was the easiest to memorize, Type II was next, and Type VI was the most difficult. As be-

TABLE 6

ACCURACY IN SORTING FROM MEMORY: EXPERIMENT III

(Percentage of *S*s correctly assigning stimuli from memory)

Problem		Perceptual Representation		
Type	Set	A (Compact)	B (Distributed: same values)	C (Distributed: different values)
I	1	96	89	100
	2	89	62	100
II	1	81	62	54
	2	96	54	42
III	1	62	50	46
	2	58	54	23
IV	1	35	50	42
	2	65	33	50
V	1	50	27	38
	2	73	42	38
VI		62	23	31

fore, the greatest accuracy on Type I was achieved with the distributed Representation C. The decrease in accuracy with increased complexity was quite gradual for the compact Representation A but was more pronounced with the two distributed Representations B and C.[10] Clearly, prior practice is not the explanation for the relationship between complexity of classification and accuracy in memorization since the results are remarkably similar to those in Experiment II. The only appreciable effect of prior experience with the task of formulating rules appears to be a reduction in the time expended in memorizing Type I classifications.

DISCUSSION OF EMPIRICAL RESULTS

We shall now try to summarize and tie together the results of the three experiments, which have just been described in detail, and to relate these results to those obtained in earlier investigations of the learning of classifications. A central feature of the three present experiments is that they attempt a systematic exploration of the effect of the *structure* of a classification upon the difficulty of learning or remembering that classification. The classifications were always constructed by dividing eight stimuli into two groups of four. Moreover, each stimulus always took on one of two highly discriminable values on each of three dimensions. The structure of each classification was defined, therefore, in terms of the way in which the membership (in one of the two classes) of each stimulus could be specified in terms of the dimensions and values of the stimuli. In

TABLE 7

TIME REQUIRED TO MEMORIZE STIMULI: EXPERIMENT III

(Median number of seconds elapsing between presentation of stimuli and *S*'s sorting of stimulus cards)

Problem		Perceptual Representation		
Type	Set	A (Compact)	B (Distributed: same values)	C (Distributed: different values)
I	1	4½	11½	12½
	2	5	14	7
II	1	9	20	23
	2	9	20	23½
III	1	12½	24	26
	2	14	25	21
IV	1	16	24	20½
	2	13	22	21
V	1	16	22½	24
	2	13	24	23½
VI		19	23½	32

[10] The differences between perceptual representations and the difference between types of classifications are both significant at the .001 level according to the chi square analysis (the use of which was described and justified for these kinds of data in Footnote 9).

addition to the structure of the classification, however, certain other features were also varied. These primarily concerned (*a*) whether the task was designed to measure learning during successive presentation of stimuli (constructed from pictures of concrete objects), or whether it was designed to measure retention after simultaneous presentation of stimuli (constructed from abstract geometrical figures); (*b*) whether the three dimensions were confined to a single compact figure, or spread out over three spatially distributed figures; and (*c*) whether *S* was confronted with a problem for the first time, or after mastering several other problems of the same kind. We turn, now, to a consideration of some of the major results of these variations and their relation to the results of earlier investigations.

Initial Difficulties of the Six Types

Figure 12 summarizes the results of the three experiments on how the structure of a classification affects its initial difficulty (i.e., when the learning or memorization of the classification was not immediately preceded by the learning or memorization of another classification of the same structural type). As can be seen, the ranking of the six structural types is the same in every case: namely, I < II < (III, IV, V) < VI (with III, IV, and V about equal in difficulty). This ranking apparently holds up, then, whether a memory or learning task is used; whether the stimuli are abstract or concrete, compact or distributed; or whether difficulty is measured by the time *S*s take to study a classification or by the number of errors they subsequently make during recall of that classification. Thus the abstract structure of a classification (as represented by the six basic types) seems to be an important determiner of its difficulty.

Many kinds of problems used in the study of concept learning resemble those originally investigated by Hull (1920) in that mastery of one of these problems can be achieved simply by discovering which of the several variable properties of the stimuli is the one that determines which response will be correct. All such problems correspond most closely to what we have termed "Type I classifications." This is true even when the relevant information is carried in a completely redundant fashion by more than one dimension of the stimuli (as in the experiment by Bourne & Haygood, 1959), since simultaneous attention to two or more dimensions is never required. However, other studies of

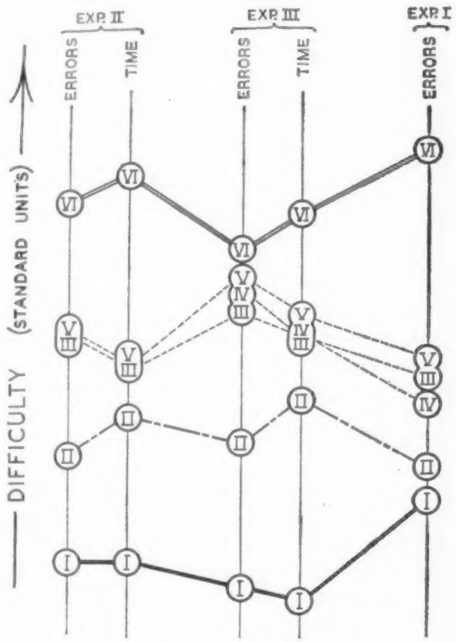

FIG. 12. Comparison of the difficulties of the six types of classifications for three different experiments and two kinds of measures of difficulty. (The data from Experiment I are restricted to the numbers of errors made during the learning of only the first problem of each type in order to make the results from this experiment comparable with those from Experiments II and III. The data from Experiments II and III are averaged over the three perceptual representations studied in those experiments—viz., A, B, C. To facilitate comparison among the five sets of difficulty scores, they were all brought into the same range by linearly transforming each set of scores so as to have the same mean and variance as the other four sets of scores. Hence the size of the units and the location of the zero points are arbitrary for each set. Finally, the results for Experiment I are placed next to the results for Experiment III so that the position of Type IV—not included in Experiment II—can more readily be compared.)

concept learning, particularly those concerned with conjunctive and disjunctive concepts (e.g., Bruner, Goodnow, & Austin, 1956; Hovland & Weiss, 1953), have come closer to our more complicated types of classifications.

Of particular interest in the present connection are two other studies, one by S. L. Smith (1954) and one by Lise Wallach (in press). The three types of classifications investigated by Wallach corresponded to our Types I, II, and VI. Wallach also used two kinds of stimuli. For the kind that most closely approximated those used in the present experiments she obtained the same ranking—i.e., (in our notation) I < II < VI. The results obtained by Wallach for her second—and rather different—kind of stimuli will be considered later. Smith's experiment differed from ours (and Wallach's) in that the stimuli varied along four (instead of three) dimensions and, consequently, in that there were 16 (instead of 8) stimuli. Nevertheless, if one of the irrelevant dimensions is disregarded, three of Smith's conditions correspond to our Types I, II, and VI. With this interpretation, his results were also consistent with ours; for again, I < II < VI. (However both Smith and Wallach stated that the differences found by them between the two more difficult types were not statistically significant.) Another condition included by Smith differed from those just mentioned in that none of the four dimensions was irrelevant. This condition was the four-dimensional analog of our three-dimensional Type VI condition and, in line with our results, was still more difficult than the three conditions just considered.

In contrast to all of these classification problems (referred to by Smith as "structured" problems), Smith (like Metzger, 1958) also included a classification problem characterized by him as "random." In this condition the 8 stimuli to be associated with one of the two responses were selected from the 16 stimuli at random rather than on the basis of the dimensions and values of the stimuli. However, our analysis of types of classifications makes clear that every classification has some structure or other. For example, if our eight stimuli were divided into two equal classes at random, the probability is 0.8 that the resulting random classification would really be a Type III, IV, or V classification. For, of the 70 distinct classifications, just 56 happen to be of one of these three types. The classification that Smith labeled random was therefore probably one of the many four-dimensional types that are roughly analogous to our three-dimensional Types III, IV, and V. (A table of these four-dimensional types has been compiled by Moore and is presented by Higonnet & Grea, 1958.) Thus the word "random" can only be interpreted as referring to the way in which the classification was generated; it cannot properly be regarded as denoting a property of the classification itself. Indeed, for every type of classification (including, therefore, any generated at random) the *S*s in the present experiments were often able to discover some kind of simplifying rule or regularity in the classification. Surprisingly, Smith's *S*s made even more errors on his random classifications than on his four-dimensional analogous of our Type VI classification.

In general, though, the present results agree with those of Smith (1954), Wallach (in press), French (1953), and others in showing that a classification is easier to learn and remember when it is related in a simple way to the dimensions and values of the stimuli. For stimuli varying along a given number of dimensions, the easiest classification is the one in which the value on a single dimension completely determines which of the two classificatory responses is appropriate. The present results, as well as those of Smith and of Wallach, show that the initial difficulty of a classification monotonically increases beyond that of this easiest classification as the values on more and more dimensions must be taken into account.[11] A similar

[11] In the present experiments there is a complete confounding of the number of relevant and the number of irrelevant dimensions such that, whenever one increased, the other necessarily decreased. This is a consequence of the facts that there were always the same number of variable dimensions (viz., three) and that the dimensions were nonredundant in the sense that every possible combination of values on these dimensions occurred with the same probability. It is only for these conditions

result was apparently also found in another unpublished experiment by Walker (referred to by Bourne & Haygood, 1959). On the other hand, the present results go beyond those that have previously been reported in that they are based on the first complete sampling of the possible types of classifications. The three new types to which attention has been called by this sampling (viz., III, IV, and V) complete the over-all pattern; they are apparently intermediate between Types II and VI both in terms of the number of dimensions that must be attended to simultaneously and in terms of the difficulty of the classification.

Effect of the Physical Representation of the Dimensions

Although the diverse kinds of stimuli used in the present experiments all led to the same ranking of the six types of classifications with respect to difficulty, they did have an appreciable effect upon the absolute level of difficulty of specific types. The most prominent difference appears to be between the distributed stimuli (Representations B and C in Experiments II and III), in which the three dimensions were presented as variations in three spatially separated figures, and the compact stimuli (Representation A), in which the three dimensions were presented as three kinds of variations in the same figure. The results show that the more difficult types of classifications (II–VI) become still more difficult when the stimuli are changed from the compact to the distributed form. Types II and VI seem to be the most strongly affected by such a change. Type I, on the other hand, is almost uninfluenced by this change. Indeed, what little effect there may be on this easiest type of classification appears to be in the opposite direction. (Note, particularly, the differences between Representations A and C for this type in Tables 6 and 7.)

The experiment by Wallach (in press) seems to have some bearing on these results. Of the two kinds of stimuli used in her experiment, one closely resembled our Representation C (in Experiments II and III). The only difference was that the two alternative figures that could occur in each of the three spatial positions of a stimulus were simple nonsense figures (composed of curved lines) rather than conventional geometrical figures. As we have already observed, Wallach's results for these stimuli were consistent with ours. The second kind of stimuli used by Wallach more closely resembled our compact Representation A in that the curved lines constituting the values on each dimension were all combined into a single, more complex nonsense figure. However, this compact representation was quite different from ours in that the values on each of the three dimensions merged into one another in such a way as to lose their identities as distinct, perceptually isolated properties. Consequently, as Wallach remarked, these stimuli tended to be reacted to as unique wholes rather than analyzed into separate dimensions and values. For these stimuli she found no significant differences in the difficulties of the three types of classifications investigated. Moreover all three were significantly more difficult than the easiest and significantly less difficult than the hardest classification with distributed stimuli. Smith (1954) also found that, when the relevant dimensions of the stimuli were made more obscure, the differences in the difficulties of different classifications tended to disappear.

---

that we propose the above generalization: namely, that the difficulty of a classification increases with the number of relevant dimensions (and, therefore, decreases with the number of irrelevant dimensions). If further dimensions were added in order to achieve independence of the number of relevant and irrelevant dimensions, the expected result would be quite different. In particular, if the number of nonredundant and relevant dimensions were held constant while further dimensions were added, the difficulty of a classification would presumably change in either of two possible ways depending upon the relevance and redundancy of these new dimensions: if the added dimensions were irrelevant to the classification, the difficulty should increase (Archer, Bourne, & Brown, 1955; Bourne & Haygood, 1959); but, if the added dimensions were completely redundant with the original dimensions and therefore relevant, the difficulty should decrease for simple classifications (Bourne & Haygood, 1959) and, apparently, increase for more complex classifications (Bricker, 1955).

In any case the present results together with those of Wallach suggest the following generalization: As the representation of the dimensions in the stimuli is made more compact, the differences in the difficulties of the various types of classifications are decreased. If the dimensions remain perceptually distinct, this compression in the variation in difficulty is primarily attributable to a disproportionate decrease in difficulty of the initially more difficult types. But, if the dimensions merge and become perceptually indistinct, the initially easier types become more difficult and, in the extreme case, all types of classifications approach the same intermediate level of difficulty.

Transfer of Classification Learning

In all three of the present experiments (as well as in that of Smith) each *S* went through several different problems in succession. However, the conditions of the various experiments differed in two respects: problems of the same basic type were either presented consecutively in clearly demarcated blocks or else intermixed at random with other problems of different types; training on each problem was either carried to a high level of mastery on each problem before proceeding to the next or else terminated after one exposure to the set of stimuli. These variations in conditions apparently influenced the extent to which the learning of one problem transferred to the next. When a high level of mastery was not required and when similar problems were scattered throughout the entire series (as in Experiments II and III), there was no systematic improvement in performance over that series. But, when a high level of mastery *was* required (as in Experiment I and the earlier experiment by Smith), fewer errors were made on the problems toward the end of the series. (However, this trend was statistically significant only in the experiment by Smith.) Finally, when both a high level of mastery was required and, also, problems of the same type were presented consecutively, the positive transfer from one problem to another of the same type was quite pronounced (as shown by the error curve in Figure 7 of Experiment I). These results appear to be consistent with the conclusions of Morrisett and Hovland (1959). They presented evidence that, in order to realize positive transfer, it is not sufficient simply to have a wide variety of problems; it is also necessary to achieve a high level of mastery on each problem. The greater amount of training insured by the learning tasks (as opposed to the memorization tasks) as well as the grouping together of problems of the same type presumably resulted in a greater mastery of the problems and problem types.

Previously reported experiments on classification learning have not usually been specifically concerned with transfer from one classification problem to another of the same type. Experiment I therefore provides new information about changes that occur in the ranking of the difficulties of the different types of classifications when several problems of the same type are learned in succession. The most striking finding, here, is that Type VI (which is initially the most difficult) also accumulates the greatest positive transfer with continued practice. As a consequence, after two or three problems in which the stimuli are changed but the type of classification remains the same, Type VI becomes less difficult than some of the other types (evidently III, IV, and V). This indicates that an exclusive focus on the initial level of difficulty of each type of classification can be misleading.

Theoretical Discussion

We now examine some of the principles that have been adduced to account for phenomena of rote learning and concept formation. The main objective will be to evaluate the ability of these principles to account for the results of the present experiments. Primary among these results is the finding that, when they are initially encountered, the six types of classifications consistently differ in difficulty according to the ranking I < II < (III, IV, V) < VI. There are also certain secondary results, though. In particular, when several problems of the same type are learned in succession, Type VI realizes by

far the greatest within-type positive transfer; and, when the relevant dimensions are distributed over spatially separated figures (rather than combined as different aspects of a single figure), the difficult types of classifications become still more difficult.

Some of the principles that have previously been proposed pertain more to the nature of the individual stimuli than to the structure of the classification and, so, do not by themselves yield definite predictions for the primary result of the present study (i.e., the ranking of the six types). It is for this reason that we omit discussion, for example, of Heidbreder's (1946, 1947) principle of degree of "thing-character" of the stimuli. In connection with this particular principle, moreover, Baum (1954) has indicated that some of the results of Heidbreder's widely-known studies may be derivable from a principle of stimulus generalization, to which we now turn.

Stimulus Generalization

A number of investigators have been attracted by the possibility that the apparently complex phenomena of concept learning might be largely understood in terms of the more elementary phenomena of rote learning. The importance of the phenomenon of stimulus generalization has been repeatedly emphasized in this regard (Baum, 1954; Buss, 1950; French, 1953; Gibson, 1940; Newman, 1956; Oseas & Underwood, 1952).[12] Baum's statement of this viewpoint is perhaps the most incisive. From the principle of stimulus generalization, as applied to verbal learning by Gibson (1940), she deduces that the difficulty of learning a classification should increase as the stimuli that are assigned to different classes are chosen to be less discriminable or as the stimuli that are assigned to the same class are chosen to be more discriminable. Certainly this principle seems to account for the obvious fact that it is easier to learn to classify four different horses as A's and four different dogs as B's than to classify two of the dogs and two of the horses as A's and the remaining two horses and two dogs as B's. For horses are more discriminable from dogs than are horses from horses or dogs from dogs. Moreover, if we assume that discriminability of two stimuli is greater when they have fewer properties in common, we can make the prediction (confirmed by our empirical results) that a Type VI classification will be more difficult to learn than a Type I. For, whereas the average number of properties shared by two stimuli that are classified together is 1.7 for Type I classifications, it is only 1.0 for Type VI (see the cubical representations in Figure 3). However, we cannot make a quantitative prediction of the difficulty of each of the six types unless we have some additional information about how the amount of generalization between two stimuli depends upon the number of properties that they have in common.

The strong interpretation of the principle of stimulus generalization for classification learning. Central to the principle of stimulus generalization is the notion that the over-all difficulty of a task is compounded primarily from the confusions of individual pairs of stimuli. From this standpoint, then, the total number of errors made during the learning of a particular classification of stimuli should be predictable from a knowledge merely of the pair-wise confusions between these stimuli. But just this knowledge can readily be obtained from experiments on identification learning: i.e., from experiments in which a different response is associated with each of the stimuli (Shepard, 1958b). Thus, if we interpret the principle of stimulus generalization to mean that the total number of times that two stimuli will be confused is the same for classification learning as for identification learning, we can predict the total number of errors for any particular classification as follows: First, several *S*s are trained to criterion on an identification task with the same set of *N* stimuli to be used in the classification task.

[12] The term "generalization" has been used to refer to various things, including an inductive inference as to what characterizes the class of stimuli to which a certain response can be appropriately extended. However, the term is used here only in the narrow sense of a primitive or automatic tendency to confuse similar stimuli during learning.

Then, the number of times each of the stimuli leads to the response assigned to each of the other stimuli is tabulated in the appropriate cell of an $N \times N$ matrix. A number in any off-diagonal cell of this matrix can be interpreted as the number of times the corresponding pair of stimuli were confused prior to reaching criterion. Now, during classification learning, the confusion of two stimuli that are assigned to the same response will not result in an overt error. Thus not all errors of identification will lead to errors of classification. In fact, in order to predict the total number of errors to be expected for any particular classification, one simply strikes out the numbers in each cell of the matrix that correspond to a pair of stimuli assigned to the same response, and sums the remaining off-diagonal entries. The predicted number of errors will in general be different for different classifications (even though the same matrix is used) because different entries are included in each sum.

The basis for this method of prediction will be referred to as "the strong interpretation of the principle of stimulus generalization in classification learning" in order to distinguish it from less quantitative formulations, such as that proposed by Baum. This "strong interpretation" is essentially an extension to classification learning of the mathematical formulation of generalization already proposed for identification learning by Shepard (1957). For, according to that earlier formulation, the confusions between stimuli resulting from stimulus generalization do not depend upon how the responses have been assigned to the stimuli.[13] There is, however, one other implication of that formulation that should be carried over to the present situation: In order to minimize the contribution of confusions between the responses to the matrix obtained during identification learning, the identification responses should be chosen to be as distinctive as possible and should be paired with the stimuli according to a different assignment for each S.

The test of the strong interpretation becomes particularly simple in cases like the present one, for which differences along each of the three dimensions of the stimuli are chosen to be about equally discriminable. We then simply determine the average number of confusions made during identification learning for pairs of stimuli with two, one, or zero values in common along the three variable dimensions. These three numbers (designated n_2, n_1, and n_0) constitute a kind of gradient of generalization. (However this gradient differs from the usual kind in that the independent variable is number of common properties rather than separation along a single physical continuum.) Now for any one of the six types of classifications, exactly 16 of the 28 possible pairs of stimuli will satisfy the condition that one stimulus of the pair is assigned to one response and the other stimulus to the other response. From an inspection of the appropriate cube in Figure 3 we can determine, for each type of classification, how many of these 16 between-class pairs have two, one, or zero properties in common. (For example, these three numbers are 4, 8, 4 for a Type I and 12, 0, 4 for a Type VI classification.) We can then calculate the total number of confusions expected for each type of classification by summing the expected number of confusions for each between-class pair. The appropriate formulas are given, in terms of the gradient (n_2, n_1, n_0) obtained from identification learning, in Table 8.

Method of testing the strong interpretation. In order to gauge the extent to which this interpretation of the principle of stimulus generalization can account for our results, we first estimated the average values of the three numbers n_2, n_1, and n_0 from the identification problems in Experiment I and, then, determined whether the substitution

[13] In the case of classification learning the following assumption is probably better: Given that two stimuli are assigned to different responses, the number of confusions between those stimuli is independent of other aspects of the assignment. Owing to the absence of differential reinforcement, the number of confusions between two stimuli might be much greater if they were assigned to the same response than if they were assigned to different responses. The possibility of a much greater number of confusions between stimuli assigned to the same response raises no problem for the present analysis, however, since such confusions are not observable anyway.

TABLE 8

FORMULAS FOR PREDICTING THE NUMBER OF ERRORS FOR EACH TYPE OF CLASSIFICATION ON THE BASIS OF STIMULUS GENERALIZATION

Type of classification	Formula for the predicted number of errors
I	$4n_2 + 8n_1 + 4n_0$
II	$8n_2 + 8n_1 + 0n_0$
III	$6n_2 + 8n_1 + 2n_0$
IV	$6n_2 + 6n_1 + 4n_0$
V	$8n_2 + 6n_1 + 2n_0$
VI	$12n_2 + 0n_1 + 4n_0$

of these three numbers into the formulas in Table 8 yielded predictions that conformed with the number of errors actually made during the learning of classifications of each of the six types. To begin with, we consider only predictions to the first classification to be learned of each type. (The case of the later problems, which is complicated by the differential within-type transfer, will be considered later.) Therefore, since the prediction is to the first problem of each type only, the numbers n_2, n_1, and n_0 were taken from the first identification problem also. Unfortunately, since the first identification problem always preceded the classification problems, this prediction might be systematically biased in the direction of overestimating the number of errors for all types of classifications. However, such an overestimation should affect all types equally and, hence, should not interfere with the prediction of the relative spacing of the six types with respect to difficulty.

Results of the test. The number of confusions between stimuli during identification learning decreased on the average as the number of properties they had in common decreased. The actual numbers, n_2, n_1, and n_0, obtained from the first identification problem were 5.03, 2.79, and 2.17, respectively. These numbers therefore conform to the kind of monotonically decreasing gradient of generalization typically found in studies of generalization during identification learning (e.g., see Shepard, 1958a, p. 246). Since generalization seems to have operated in the expected manner in the identification problem, then, the conditions are appropriate for the test of whether the prediction to the classification problems is also successful. In Figure 13 the number of errors actually made on the first classification problem of each type is plotted against the number of errors predicted from the previously obtained gradient, n_2, n_1, n_0 (and the formulas in Table 8). In contrast to the agreement with our expectations for identification learning, the prediction to classification learning clearly failed. The predicted numbers of errors were too great for all except perhaps Type VI; the amount of variation between the predicted numbers of errors was strikingly smaller than the amount of variation between the actual numbers; and, finally, the predicted ranking of the difficulties of the six types was itself incorrect. (Note that, although I and VI were correctly predicted to be the easiest and most difficult classifications, II was erroneously predicted to be next to VI rather than next to I in difficulty.) The fact that the first identification problem always preceded the

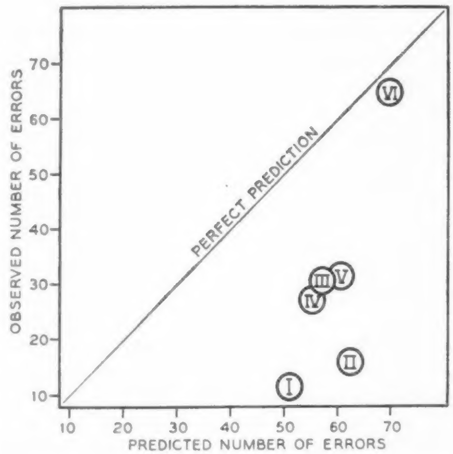

FIG. 13. Mean number of errors made when each type of classification was learned for the first time, plotted against the number of errors predicted from the gradient (n_2,n_1,n_0) obtained during identification learning. (The departure of the six points from the 45-degree line represents a predictive failure of the strong interpretation of the principle of stimulus generalization in classification learning.)

first classification problem might in part account for the first kind of failure of the prediction, but it presumably could not account for the remaining two. The results of the test seem clear then; the strong interpretation of the principle of generalization cannot by itself account for the difficulties of the different types of classifications.

Tests of a weaker interpretation. Of course one could object that the strong interpretation of the principle of generalization is *too* stringent. In particular, one might argue that the gradient of generalization is not fixed but, rather, changes in some systematic way when the experiment is converted from one on identification learning to one on classification learning. However, contrary to this argument it can be demonstrated that every possible gradient that might be assumed leads to an incorrect prediction. The demonstration proceeds as follows: First, since the gradient consists of just three numbers (n_2, n_1, n_0), any possible gradient is uniquely representable as a point in the +++ octant of the three-dimensional Euclidean space with Cartesian coordinates n_2, n_1, and n_0. Moreover, according to the formulas in Table 8, multiplication of the three numbers of a gradient by the same constant affects the prediction only of the absolute number of errors, but not the relative spacing among the six types. For purposes of predicting only the relative spacing, then, we can restrict consideration to the set of normalized gradients for which $n_2 + n_1 + n_0 = 1$. Each such gradient is represented by a point on the triangular plane with vertices at the points (1, 0, 0), (0, 1, 0), and (0, 0, 1) as illustrated in Figure 14. The triangular region is partitioned in the figure to show what the general shape of the gradient is for various subregions of the total triangle. (For example, all monotonically decreasing gradients can be seen to fall in two triangular sectors on the left of the total triangle.) The point of intersection of all the triangular sectors at the center corresponds to the flat gradient with $n_2 = n_1 = n_0 = 1/3$.

Now to each point in the triangular space of normalized gradients there corresponds a prediction of the relative difficulty of each of the six types that can be directly determined simply by substituting the three numbers for that point (n_2, n_1, n_0) into the six formulas in Table 8. This space can therefore be systematically explored to see whether any gradient exists that yields the correct ranking of the six types. Figure 15 summarizes the results of this exploration. (The triangle exhibited there is the same as the one illustrated in Figure 14 but, for convenience, is now presented as normal to the line of regard.) This triangular space is partitioned into sectors within which the same ranking holds (although the relative spacing of the types changes continuously from one point to another within any sector). For gradients falling on any boundary line separating two adjacent sectors, the predicted ranking contains a tie. Such a tie is always between those types that change rank orders in moving from one sector to the other. All six types are therefore tied at the central point of intersection of all boundaries.

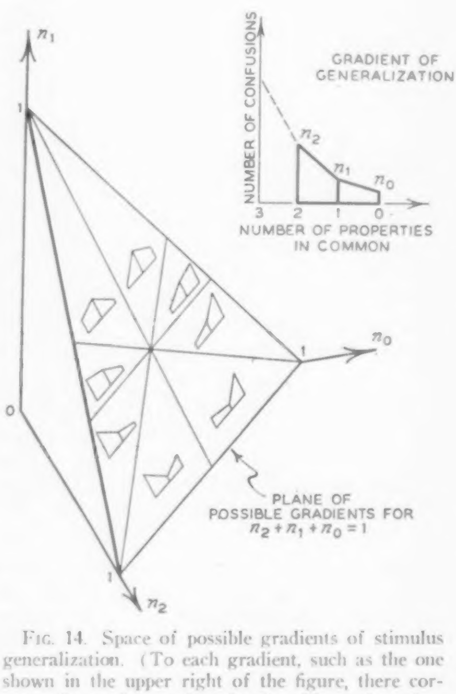

FIG. 14. Space of possible gradients of stimulus generalization. (To each gradient, such as the one shown in the upper right of the figure, there corresponds a point on the triangle with vertices situated one unit out on the three orthogonal axes.)

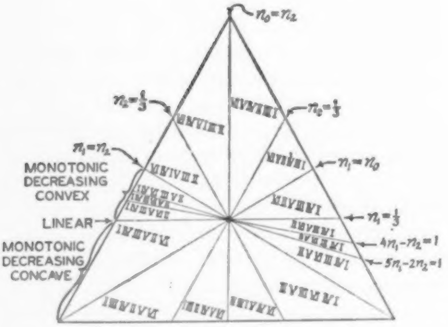

FIG. 15. Ranking of the difficulties of the six types of classifications predicted for every possible gradient of generalization. (Equations are given in terms of the Coordinates n_2, n_1, and n_0 for the lines that divide the triangular space of possible gradients into regions for which the same ranking holds.)

The first thing to note is that, although there are $(6!)2^5$ or 23,040 possible rankings of the six types (including ties), only 16 + 16 + 1 or 33 of these can be generated by varying the shape of the gradient of generalization. There is therefore a real question as to whether there exists a gradient that will make the correct prediction. This question must be answered negatively, however, for at no point in the triangle does the ranking I < II < (III, IV, V) < VI occur. The only close approximation is near the bottom of the central vertical boundary where the ranking is (I, II, III) < (IV, V) < VI. But this ranking clearly departs from the empirical pattern exhibited in Figure 12. Moreover it requires an implausible gradient in which stimuli that do not have the same value on any of the three dimensions are confused more frequently than stimuli that do have one common value (see Figure 14).

The attempt to account for performance in a classification task in terms of the principle of generalization alone also fails in another way. For, although this principle is not necessarily inconsistent with a general improvement in performance on successive problems, it does not seem capable of accounting for any variations (from one type to another) in the amount of within-type transfer. Thus, not only the initial ranking, but also the subsequent shift (shown in Figure 6) to the different ranking with VI < (III, IV, V) is inexplicable on the basis of this principle.

One reason for the failure of the generalization principle. The fact that the strong interpretation of the principle of stimulus generalization yielded a reasonable prediction of difficulty only for Type VI (see Figure 13) suggests that the most serious shortcoming of the generalization theory is that it does not provide for a process of abstraction (or selective attention). The argument runs as follows: In a Type I classification *S* notices that the values on one of the three dimensions are highly correlated with the classificatory responses. That one dimension then becomes the focus of *S*'s attention. The stimuli of the between-class pairs will still have properties in common; but, since all of these shared properties are on the now unattended dimensions, they will no longer mediate generalization to the same extent as in identification learning (where, in order to respond correctly, *S* must attend to all three dimensions). By abstracting the relevant dimension, then, *S* might keep the total number of errors in a Type I classification well below that predicted from the generalization theory. A similar argument can be developed for Type II and, in a slightly modified form, for Types III, IV, and V. In Type VI, however, there is no opportunity for abstraction in this sense; for, in order to respond correctly in a Type VI problem, *S* must take account of all three dimensions (just as in identification learning). Figure 13 suggests that, when this kind of abstraction is precluded, the generalization theory alone may account for the initial difficulty of a classification. Further support for this distinction between generalization (or stimulus confusion) and abstraction (or selective attention) will be presented when we come to the discussion of individual differences. We shall also argue that the marked positive transfer observed within a series of Type VI problems is evidence for a somewhat different kind of abstractive process. Meanwhile, however, we need to examine some notions that might be thought to account for the simple abstraction of—or

selective attention to—relevant dimensions of the stimuli (cf. also Binder & Feldman, 1960, pp. 15–22).

Conditioning of Cues

Following the explanatory successes of certain theoretical notions introduced particularly by Estes (1950, 1959b), many learning theorists are now predisposed to regard a stimulus as a collection of elements or cues each of which can separately become conditioned to a response. The application of this idea to the learning of classifications seems at first rather straightforward. In a given Type I classification, for example, Response A might always be reinforced in the presence of the cue black but never in the presence of the cue white. Conversely Response B would always be reinforced in the presence of the cue white but never in the presence of the cue black. However both responses would be reinforced just half of the time in the presence of each of the other cues: large, small, triangular, circular. In this way we are apparently provided with an account of how Response A comes to be associated with the black stimuli and Response B with the white stimuli regardless of their size and shape.

Furthermore, as pointed out by Bush and Mosteller (1951), this kind of theory might even subsume the principle of stimulus generalization. In particular, since the probability that a response will be made to a given stimulus is generally assumed to be equal to the fraction of the cues (in the given stimulus) that has been conditioned to that response, the probability of a response that has been conditioned to a particular stimulus should fall off linearly for stimuli that have two, one, and zero properties in common with that stimulus. And, as we shall indicate later, although most of the generalization gradients actually obtained from the identification condition were somewhat concave upward, many were nearly linear.

A closer examination, though, reveals that the performance of *S*s who behaved in accordance with this theory could never approach the degree of accuracy that is empirically observed. In the example considered above (in which only the cues of color are relevant to the classification) the presence of size and shape cues, half of which are always conditioned to the wrong response, must influence *S*s to respond incorrectly to a substantial fraction of the presentations, regardless of how long training might be continued. For the same reason, *S*s would never be able to reach criterion on an identification problem in which the stimuli consisted of overlapping collections of cues. We now consider two elaborations of the basic cue-conditioning idea that have recently been proposed in attempts to correct this deficiency; namely, the pattern model of Estes and the adaptation model of Restle.

Conditioning of patterns of cues. The pattern model described by Estes (1957, 1959a, 1960) specifies that responses can become connected not only to the individual cues as independent elements of the stimulus but also to the total pattern of cues that uniquely constitutes each stimulus. Since, as noted above, the asymptotic performance of actual *S*s surpasses that predicted by the original cue-conditioning model, Estes (1957, p. 616; 1960, p. 60) concludes that the total patterns must eventually prevail in controlling the responses. On the other hand, there is some evidence that the conditioning of individual cues predominates during the early phases of learning (Estes, 1957). This may in part be attributable to the fact that the cues are more frequently available than the patterns. Thus in our experiments a single cue (e.g., the color black) occurs on half of the presentations, whereas a single pattern (e.g., the large black triangle) occurs on only an eighth of the presentations. In any case, if the responses eventually come under the exclusive control of the patterns, the performance of *S*s will approach 100% correct on either identification or classification problems (as actually observed).

Several difficulties still remain, however. First, the admission that patterns of cues can themselves become directly connected to responses removes some of the appeal of the cue-conditioning model. The original model was rather close to a description of a physically realizable mechanism. Indeed Rosen-

blatt's "Perceptron" (1958) might even be regarded as one way of physically realizing the kind of general idea formalized in Estes' original model. The pattern model, although perfectly permissible as a formalism yielding testable predictions, seems to leave more of the inner mechanics unspecified. In particular, the details of the process whereby the individual cues become fastened together into a functional unit that can be directly attached to a response remain mysterious.

Of course one could argue that, since the pattern can enter into a unitary relation with a response, it is itself just another cue that was part of the stimulus all along. The only unique feature of this particular cue is that (unlike the others) it is not shared by any other stimulus. Such an argument would, indeed, be completely consistent with Estes' (1960, p. 52) definition of cue or "stimulus element." Unfortunately, though, it points up yet further problems. In order to master an identification or classification problem, the conditioning of the cue corresponding to the total pattern of other (component) cues must completely override the connections (previously formed) between these component cues and the responses. But no specific rules seem to have been given by Estes for the process of eliminating these earlier connections. Finally, if the attainment of criterion is possible only because the total patterns become conditioned to their appropriate responses, then (even though the initial rate of learning might vary from one classification to another) the final mastery of the different types of classifications would presumably be achieved after about the same number of trials. Thus, in order to account for the rapidity with which actual *S*s reach criterion on a Type I classification, the pattern model (like the generalization model) seems to require the annexation of an additional mechanism for the selective suppression of irrelevant cues.

Adaptation of cues. One such possible suppression mechanism has been proposed by Restle (1955, 1957). Restle's idea is that cues that are uncorrelated with the reinforcement of the responses become "adapted" and, hence, lose their control over those responses. Thus, in the classification problem considered before (in which only the color of the stimuli is relevant), the cues of size and shape would adapt out leaving the black and white cues in complete control of the responses. With this additional principle, then, a model based upon the independent conditioning of the cues of a stimulus seems to provide a mechanism for abstraction such as our discussions of stimulus generalization and pattern conditioning led us to seek. Indeed, Bourne and Restle (1959) have recently shown that a variety of phenomena of concept learning can be accounted for by a model of this kind.

Unfortunately, in order to account for the mastery of an identification problem (or, indeed, of a Type VI classification problem), certain additional complications of the model are necessary. For example, in an identification problem, since a different response must be associated with each of the eight stimuli, all confusions must eventually be eliminated between stimuli that differ in color. But, as we have seen, this is possible only if the cues of size and shape become adapted. This, in turn, would preclude the elimination of confusions between stimuli differing only in size and shape. Bourne and Restle, in their discussion of a four-response problem, apparently cope with this difficulty by considering, in effect, that cues do not become adapted absolutely but only in relation to the pairs of responses for which those cues are irrelevant. Such a complication of the notion of adaptation seems to us to decrease its attractiveness as an account of the phenomenon of abstraction or selective attention. Furthermore, although the model of Restle and Bourne specified how a cue that is known by *S* to be irrelevant becomes adapted, it does not specify how *S* comes to know that a cue is irrelevant.

A general dilemma faced by cue-conditioning models. Beyond the specific objections raised against the models of Restle and Bourne and of Estes, these as well as other models for the conditioning of cues face the following more general difficulty: On the one hand, the cues might be identified with the elementary physical properties of the stimuli —e.g., largeness, smallness, blackness, whiteness, etc. (This is what Bourne and Restle

appear to do in their discussion of two-response experiments.) But then we should have to predict that the performance of *S*s on a Type VI classification would never improve beyond its initial chance level, for none of the elementary properties in this type of classification is by itself correlated with the reinforcement of either response. On the other hand, we might consider that not only the elementary properties but also any pattern of these elementary properties to which *S*s can learn to respond differentially can serve as a cue. (This seems to be the original intention of Restle, 1955, pp. 11, 18; and of Estes, 1960, p. 52.) But under this interpretation we are left with the problem of specifying, for each possible pattern of cues, some parameter (e.g., a weight for that pattern) governing the rate at which it can become conditioned to a response. This, in turn, reduces to our original problem; namely, the problem of determining the difficulty of each possible classification.[14]

Thus, although a theory based upon the notions of conditioning and, perhaps, the adaptation of cues at first showed promise of accounting both for stimulus generalization and abstraction, further investigation indicated that it does not, in any of the forms yet proposed, yield a prediction of the difficulty of each of our six types of classifications. Nor does this kind of theory seem to account for the relatively much greater positive transfer found in the case of Type VI classifications.

Abstraction and the Formulation of Rules

As we have just seen, the hypothesis that responses can be conditioned only to elementary properties (i.e., to the properties that define Type I classifications) seems to be disconfirmed by the fact that *S*s can learn Type VI classifications. On the other hand, the hypothesis that responses can be conditioned to arbitrary combinations of elementary properties (such as that defining a Type VI classification) removes the predictive force of the conditioning models. Moreover, the notion that such a combination of elementary properties can itself serve directly as a cue seems implausible in view of the kinds of verbalizations actually produced by *S*s. In describing a single stimulus, our *S*s used words like "large," "black," etc. (in Experiment II) or like "candle," "trumpet," etc. (in Experiment I). These clearly referred to the elementary properties. For the statement that a stimulus is "black," for example, is essentially equivalent to the statement that it belongs within the group of black stimuli that are set apart by the Type I classification based upon color. In no instance was an individual stimulus described by referring to a property that would correspond in this way to the particular group of stimuli set apart by, say, a Type VI classification. Indeed, even when the stimuli are sorted out according to a Type VI classification, *S*s are unable to see the four stimuli in either class as having any one property in common. And when (as in Experiment I) some *S*s eventually do discover a way of characterizing the stimuli that go together in such a classification, they invariably do this by formulating an elaborate rule in terms of the elementary properties. As an example of a relatively simple rule, they might finally say: "The figures on the left must be black and small and triangular or else have just one of these three [Type I] properties." But even after discovering a rule of this kind, *S*s do not then regard it as a unitary property of a stimulus and, surely, they would not subsequently invoke it in describing a single stimulus—however completely.

Outline of an alternative to the conditioning models. In view of the preceding considerations, we are led to consider that only the properties defining Type I classifications act directly as cues, and that classifications other than Type I can be learned only by con-

[14] Dattman and Israel (1951) have proposed a principle to account for Heidbreder's results that is very similar to the notion that the cue corresponding to each possible classification has a certain weight (or "salience"). According to them: "the relative ease with which concepts are attained is directly dependent upon the degree of perceptual effectiveness with which the instances serve to present the features to be conceptualized." This principle evidently suffers from the same lack of predictive force as the weighted-cue idea, since no objective method is proposed whereby the "perceptual effectiveness" can be measured independently of the results of the concept task itself.

structing appropriate rules for them in terms of Type I classifications. Accordingly, *S*s are no longer regarded as passively confronting one population of cues after another while a certain crucial subset becomes gradually connected to the correct response. Rather, they are regarded as actively abstracting (or attending to) dimensions, and then formulating and testing rules about how the values on those dimensions combine and interact to determine which classificatory response will be correct. The development of an explicitly detailed model for this kind of process will not be undertaken here. The present study, together with others, might serve as a useful basis for such an undertaking. But the development itself would require a further investigation implemented, perhaps, by the new tool of computer simulation (Hovland & Hunt, 1960; Newell, Shaw, & Simon, 1957; Newell & Simon, 1959). Nevertheless the tentative description of the learning process, given here only in general outline, is sufficient to lead to certain expectations about the relative difficulties of our six types of classifications as well as about how these difficulties are influenced by certain of the conditions imposed in the present experiments.

The initial difficulties of the six types. If the foregoing description of the learning process is correct, the difficulty of any given classification should be directly related to the complexity of the rules required to build it up out of Type I classifications. That these rules may be largely formulated and used at the verbal level is suggested by the high correlations, found in both Experiments I and II, between accuracy of performance (in sorting or responding to stimuli) for each type of classification and the simplicity of the rules that could be stated by *S*s for the same type of classification. Of course the rules verbalized by the *S*s varied in detail from one *S* or occasion to another. Still, since the number of possible combinations of values increases very rapidly with the number of dimensions considered, the difficulty of a classification should increase with the number of dimensions (or Type I classifications) that are required for the specification of that classification. Thus we should at least expect the ranking I $<$ II $<$ VI. Furthermore, the other types (III, IV, and V) are presumably intermediate between II and VI. For, although all three dimensions are relevant for each of these classifications, some (but not all eight) of the stimuli can be properly classified by knowing the values on just two of the three dimensions.

Actually there are many different measures that could be used to express the difficulty of building up a classification out of Type I classifications. We could, for example, use the minimum length of a logic expression that defines the classification by means of conjunctions and disjunctions of the elementary properties. (We already noted in the introduction that such an expression is much longer for a Type VI than for a Type I classification.) We could also base it upon the number of relay contacts required for the physical realization of the Boolean function corresponding to the given classification (Higonnet & Grea, 1958). Or we could base it upon the average number of single-dimensional binary decisions required to place a randomly selected stimulus in the appropriate class when the sequence of decisions is made in the optimum order. But these (and many other) measures all agree with the ranking derived above in a more informal manner. In particular, they are all consistent with the ranking I $<$ II $\leq$ III $\leq$ IV $\leq$ V $<$ VI. The only disagreements concern the predictions of ties among the intermediate types (II through V). The one quantitatively defined measure that has seemed most satisfactory to us reflects the extent to which the information about the classification is distributed over the three dimensions (rather than confined to a single dimension—as in a Type I classification). In the appendix this measure is shown to lead to the ranking that corresponds most precisely with that found empirically; namely, I $<$ II $<$ (III, IV, V) $<$ VI.

Transfer of classification learning. The tentative description of the learning process in terms of abstraction and the formulation of rules also has some implications for the transfer of classification learning. In particular, if *S* is faced with a new classification problem that, however, he has reason to

believe is of the same type as several preceding problems, there is the possibility of a certain amount of positive transfer. If, for example, the problems have all been of Type II, *S* might learn to proceed directly to testing rules about the interaction of values on pairs of dimensions without wasting any time in testing useless one-dimensional rules.

Actually the most pronounced positive transfer was observed in the case of Type VI, and the unique degree of transfer in this case probably comes from the intervention of a somewhat different process. In particular, when *S* has formulated the odd-even rule, he has apparently performed an abstraction on a higher level than in the case of the simple abstraction of relevant dimensions. Indeed this higher level abstraction is most effective precisely when the abstraction of individual dimensions would preclude solution of the problem—i.e., for Type VI. The effectiveness of the odd-even rule can be appreciated by considering that *S*s who apply it with maximum efficiency need only learn the response to the first stimulus; each following stimulus then calls for that same response if and only if it has an odd number of properties in common with the first stimulus. Indeed, the one *S* in Experiment I who most completely mastered this odd-even rule (viz., S_4) made altogether only one error on the last four Type VI problems. This is to be compared with the 50 errors she made during the first Type VI problem (before discovering this powerful reductive rule).

This kind of consideration can be used to predict the ranking of the asymptotic difficulties of the six types: i.e., the ranking that presumably would be achieved if a sufficient number of problems of the same type were consecutively supplied. For, if *S* has really abstracted a rule that uniquely defines a given type of classification, the difficulty of learning a new classification of that type should be directly determined by the fraction of the 70 possible classifications that corresponds to that type. Thus Type VI should become even easier than Type I (as it did for S_4) because, whereas six of the 70 possible classifications are Type I, only two are Type VI. If *S* knows that he is going to have a Type I problem, he knows that only one dimension will be relevant, but he does not know which of the three dimensions this is and he does not know which value on this dimension goes with each of the two responses (hence the 3×2 or six possibilities). But, if *S* knows that he is going to have a Type VI problem, his only uncertainty concerns which response goes with the first stimulus (hence the two possibilities). Now the number of classifications or possibilities for each of the six types (I, II, III, IV, V, and VI) are, respectively, 6, 6, 24, 8, 24, and 2. Therefore the predicted asymptotic ranking of these types is VI < (I, II) < IV < (III, V). The most striking difference between this ranking and the consistently found initial ranking is the change in the position of Type VI from the extreme of greatest difficulty to the extreme of least difficulty. The fact that our *S*s did not all achieve the predicted asymptotic ranking can be taken as an indication that the five consecutive problems of the same type did not provide enough opportunity for all *S*s to discover a rule (like the odd-even rule) that is sufficiently abstract to carry over to a new set of stimuli. However, the marked drop in errors over the five consecutive Type VI problems (Figure 6) supports this analysis of transfer of classification learning.[15]

---

[15] Although the odd-even rule for Type VI greatly simplified the task of learning such a classification, this rule is not easy for naive *S*s to discover. Most *S*s reach criterion on the first few Type VI problems by means of much less efficient rules. Thus, *S*s who for some reason had previously had considerable experience with Type VI classifications might not find Type VI so difficult initially.

It might also be remarked here that there is probably a connection, in general, between the reductive power of the best rule for a type of classification and the number of distinct classifications that are of that type. The number of classifications of a given type has to do with the "symmetry" of that type: i.e., with the number of transformations (rotations and reflections) of the cube in Figure 2 that do not change the type of the classification (Higonnet & Grea, 1959). And symmetry seems to be the basis of reductive rules. For example, the uniquely powerful odd-even rule for Type VI is made possible by the fact that, in that type alone,

Effect of the physical representation of the dimensions. We now consider why the spacing between the difficulties of the types is affected by whether the dimensions are represented in a single compact figure or in three spatially distributed figures. There seem to be two related possibilities: one emphasizing the verbal and the other the perceptual aspects of the classification task.

The verbal counterpart for a stimulus in the compact representation (A in Experiments II and III) might be, simply, "large black triangle." The counterpart for a stimulus in the distributed Representation C would be "large circle, white triangle, shaded square." As far as a Type I classification is concerned, this would not be expected to make much difference in the complexity of the rule. The rules might be: "If the figure is black, put the card on the left" (for A) vs. "If there is a shaded figure, put the card on the left" (for C). These two rules seem about equally complex. On the other hand, the typical kinds of rules given for the more difficult classifications tend to be longer for Representation C. Thus, for Type V, the two rules might be: "Black figures go on the left and white on the right except for the large black triangle which goes on the right and the large white triangle which goes on the left" (for A) vs. "Those containing a black circle go on the left and those with a white circle go on the right, except that if there is a large triangle, black circle, and shaded square it goes on the right and if there is a large triangle, white circle, and shaded square it goes on the left." This difference in the lengths of the verbal rules might account, in part, for the greater difficulty of the complex classifications with the distributed representations.

The other explanation is somewhat different. It argues that, in the case of the distributed representations (Experiment I or B and C in Experience II and III), *S*s can directly perceive only the elementary properties of the stimuli. As already proposed, then, they would have to build up each classification on the basis of Type I classifications alone. For the compact Representation A, however, it is possible that some of the simple interactions between dimensions are perceived more or less in the same way as elementary properties. For example, a large black triangle might be immediately seen as a black triangle without having to construct this fact out of the two component facts that it is black and that it is triangular. In the case of the compact representation, then, Estes' notion that patterns of elementary cues can themselves serve as cues becomes rather plausible—at least for these simple "conjunctive" patterns. The implication for the present experiments would primarily be a decrease in the difficulties of the initially more difficult types when compact stimuli are used. For these types would no longer have to be built up exclusively from the elementary properties; they could now utilize simple conjunctions of these as well. Thus, in a Type III classification, *S*s would not have to learn to group all four stimuli in one class together (e.g., large black triangle, large black circle, small black circle, and small white circle) but, rather, they would only have to learn to group two kinds of stimuli in that class (e.g., large black figures and small circles). Such an explanation might also account for the slightly greater difficulty of the compact representation for Type I noted in the discussion of the empirical results. For, whereas the spatial separation of the dimensions might help *S*s to isolate the single relevant property, the addition of conjunctive properties through the compact representation in effect confronts *S*s with a greater number of properties from which the correct one must be selected. This conjecture is supported by the relatively large number of two-factor rules formulated for Representation A of Type I (see Table 3).

Individual differences. Finally, the distinction that we have made between generalization and abstraction suggests that a given classification could in principle be learned either by rote or by concept. If for

each of the three dimensions plays the same role. This is most clearly shown in Representation B of Experiments II and III. Here one need only know that either one or all three of the component figures are large to know that the stimulus belongs to a particular class; it is not necessary to know just which of the three figures these are.

some reason S did not abstract or selectively attend to relevant dimensions or formulate reductive rules, the difficulty of a classification would presumably be predictable from the principle of generalization: i.e., from the confusions made during identification learning. In this case, since identification learning is essentially a rote task, we might reasonably assert that the classification, also, was learned by rote. If, on the other hand, S abstracted the relevant dimensions and formulated a rule specifying how the values on these dimensions determine the response, the difficulty would no longer be predictable from identification learning alone. In this case S might be said to have proceeded in a conceptual manner. One source of evidence for these notions comes from an examination of certain differences in the performances of individual Ss.

Figure 16 shows again the triangular space of possible gradients of generalization previously presented in Figures 14 and 15. The points denoted by the circles labeled 1 through 6 are each based upon the average of the gradients (n_2, n_1, n_0) obtained from the five identification problems for each of the corresponding six Ss (S_1–S_6) in Experiment I. The point denoted by the circle labeled x is based upon the average of the gradients for eight additional Ss who were run on one identification problem each in order to secure further data about generalization. With one exception, these average gradients are confined to a small region of

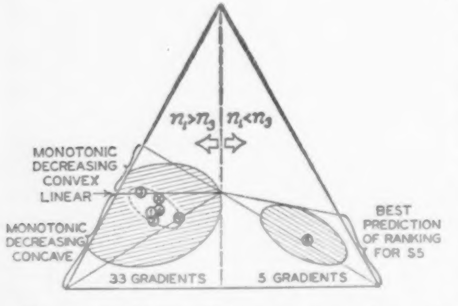

FIG. 16. Individual differences in gradients of stimulus generalization. (Averages of gradients actually obtained during identification learning are plotted, for different Ss, in the triangular space of possible gradients.)

the total space of possible gradients and, indeed, are monotonically decreasing and either concave upwards or nearly linear. Moreover, all 33 of the individual gradients upon which these six similar average gradients are based fall in the larger shaded region in the left of the total triangle. For none of these 13 Ss did the strong interpretation of the principle of generalization predict the obtained ranking of the types of classifications with respect to difficulty.

One of the Ss (viz., S_5), however, consistently differed from the others. As can be seen, her five individual gradients all fall (along with their average) within the shaded region on the right of the total triangle. Strikingly, there is no overlap between the two shaded regions. Whereas for every single gradient of the other 13 Ss, $n_2 > n_0$, for all five of S_5's gradients, $n_2 < n_0$. That is, this S invariably confused stimuli having none of the three variable properties in common more frequently than stimuli having two of these properties in common. Stranger still, this S made more errors on the "easiest" type of classification (Type I) than on any other type.

An examination of the rules verbalized by this S seems to provide an explanation for these puzzling reversals. Throughout the series of 25 problems with nonoverlapping stimuli, this S consistently employed a particular recoding scheme—one used only in rare instances by other Ss. In applying this scheme, S_5 would begin each problem by arbitrarily picking two of the eight stimuli having no properties in common. The three pictures constituting one of these two anchoring stimuli would arbitrarily be called Group 1 and the three pictures constituting the other, Group 2. Each of the remaining six stimuli was then specified in terms of the two anchoring stimuli by stating whether a Group 1 or Group 2 picture appeared in each of the three possible positions. In this way the eight stimuli shown in Figure 4 were recoded into the eight patterns exhibited in Figure 17. The employment of this recoding system means that S_5 necessarily always took account of all three dimensions of each stimulus and, hence, forfeited the possibility of abstracting the single

relevant dimension in the Type I classifications. For example, the four stimuli on the left in Figure 17, which originally shared the same property (viz., the candle in the lower left position), no longer seem to have much in common. This, then, might help to account for the large number of errors that S_5 made on Type I problems.

In addition, though, this recoding scheme should have a pronounced effect on the gradient of generalization. The argument runs as follows: an examination of the errors most frequently made during the learning of Morse code (Keller & Taubman, 1943; Plotkin, 1943) reveals that errors in which a single element is mistaken (e.g., a dot is taken for a dash) are less common than errors in which the entire pattern of dots and dashes is transformed as a whole. The most frequent errors, indeed, seem to involve reversals (in which the temporal pattern is confused with the same pattern taken in reverse order—e.g., – – • for • – –) or complementations (in which all the dashes are converted to dots and vice versa—e.g., – – – for • • •). Similarly, then, the most frequent confusions to be expected between the patterns in Figure 17 are not those that are produced by altering a single element (e.g., by changing a 1 to a 2) but those that are produced by transforming the pattern as a whole. In particular, confusions should be most common in which every 1 is converted into a 2 and every 2 into a 1 and, perhaps to some extent, in which the pattern suffers a left-right reversal. Now, if each pattern in Figure 17 is most frequently confused with its complement (i.e., the one that is adjacent to it on the right or left), we are provided with an explanation for the unusual gradient in which $n_0 > n_2$; for just these pairs of patterns correspond to the pairs of stimuli that have no properties in common. By extending this argument in detail, a plausible case can be made for the statement that Ss who used this recoding scheme should also have $n_0 > n_1 = n_2$: i.e., their gradients should fall on the line connecting the center of the triangle and the lower right corner. As can be seen in Figure 16, most of the gradients obtained from S_5 do in fact fall in the vicinity of this line.[16]

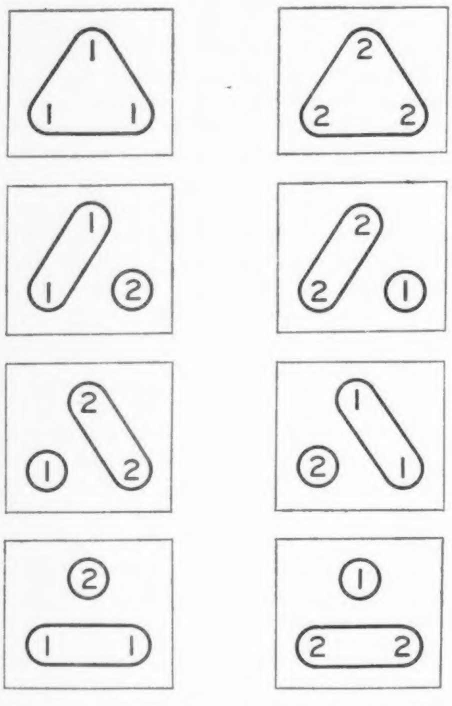

FIG. 17. The eight stimuli of Figure 4 recoded according to the system used by S_5. (The classification shown here is the same as that shown in Figure 4; but the order of the four stimuli on the right has been changed so that the stimuli that are most similar after the recoding are horizontally adjacent.)

Now, since S_5's coding of the stimuli prevented her from abstracting the relevant dimensions, we should expect the strong inter-

[16] As already mentioned, we have used "generalization" to refer to confusions between stimuli. But, in accordance with the above discussion, we should now acknowledge further that these confusions might arise from at least two possible sources: They might be the result of "primary" stimulus generalization based upon the similarity of the stimuli in terms of their physical properties (e.g., the number of such properties that they have in common). This assumption seems to be consistent with the gradients that we observed for all Ss except S_5. Or, on the other hand, these confusions might be the result of "mediated" generalization based upon the similarites of the implicit (recoding) responses made to these stimuli by a particular S. This is what we have proposed for S_5.

pretation of the principle of generalization to predict the initial difficulties of the different types of classifications. Of the four types of classifications administered to S_5, the ranking with respect to difficulty was found to be (V, II) < VI < I. (The number of errors on the first problem of each of these types was, respectively, 28, 29, 38, and 55.) The predicted ranking for the triangular subregion containing S_5's average gradient (and most of the surrounding shaded area) is II < V < VI < I. The only discrepancy between these two rankings is in the case of II and V. And this discrepancy may be attributable to the fact that S_5 had Type II first (see Table 1); for, as shown in Figure 5, *S*s tended to make more errors on the very first classification problem. The rough agreement between the predicted and the obtained ranking of problem difficulties for this *S* is emphasized by the fact that they both depart strikingly from the predicted and obtained rankings for all other *S*s.

Summary

A combined empirical and theoretical investigation of the difficulties of different kinds of classifications was undertaken using both learning and memory tasks. Sets of stimuli of a variety of kinds were used but, in each set, there were eight stimuli each of which took on one of two possible values on each of three different dimensions. For example, in one set, each stimulus was large or small, black or white, and triangular or circular. The classifications to be learned or remembered were always set up by assigning four of the eight stimuli to one class and the remaining four stimuli to the other class. Three kinds of procedures were used. In one, *S*s learned to associate one of two classificatory responses (e.g., A or B) with each of the eight stimuli by means of a method of successive presentation and response correction (the usual paired-associate procedure). In the two other procedures, *S*s were presented with a simultaneous array of the eight stimuli already grouped into the two classes and then, after the removal of the array, were tested for their ability either to sort the stimuli into the same two classes or else to state a concise rule specifying how the stimuli were classified in terms of their dimensions and values. These procedures were used to determine the difficulties of all possible types of classifications of the stimuli into two groups of four. In addition, transfer of classification learning and the effect of representing the dimensions of the stimuli in different ways were also investigated. Finally, various mechanisms that have been proposed to account for phenomena of rote learning and concept formation were evaluated in relation to the empirical results. The following conclusions were drawn:

1. Of the 70 possible classifications of the eight stimuli into two equal groups, there are only six basic types. The different classifications belonging to any one of these types have the same structure; they differ only with respect to which of the three dimensions is assigned to which of the three roles in the classification, and with respect to which of the two classificatory responses is assigned to which group of four stimuli. These six types we denoted by the roman numerals I–VI. For the purposes of the classification, only one dimension is relevant for Type I, two for Type II, and all three for Types III–VI. These last four types differ, however, in the ways the values on the three dimensions interact in defining the classification.

2. When classifications of these six types are encountered for the first time, they consistently differ in difficulty according to the ranking I < II < (III, IV, V) < VI (with III, IV, and V about equal in difficulty). The same ranking is found for learning and memory tasks, inspection time and error scores, and a variety of different kinds of stimuli.

3. When several classifications of the same type (but each using a different set of stimuli) are learned in succession, the different types of classifications accumulate differing amounts of positive transfer. As a consequence the initial ranking of the difficulties of the six types changes so that, after several consecutive problems of the same

type, Type VI (which is initially the most difficult) becomes easier than some of the other types.

4. When the stimuli are changed so that the three dimensions are represented as variations in three spatially separated figures rather than as three kinds of variations in a single compact figure, the ranking remains the same but the spacing between the difficulties increases. In particular, the more difficult types of classifications (II–VI) become still more difficult, but the easiest (Type I) remains about the same or even decreases slightly in difficulty.

5. A high correlation exists, over the various conditions of the experiments, between the performance measures (in terms of time or error scores) and the simplicity of the verbal rules that *S*s can formulate to describe the classifications.

6. The cue-conditioning models, including the recent "pattern" and "adaption" models, do not seem to yield predictions of the difficulties of the six types of classifications.

7. The principle of stimulus generalization, on the other hand, can be interpreted in a form that does yield testable predictions for the learning of classifications. The predictions generated by the particular interpretation attempted here, however, do not agree with the empirically determined difficulties of the six types. Apparently the principle of stimulus generalization does not by itself provide an account for the fact that *S*s can abstract the relevant from the irrelevant dimensions of the stimuli.

8. The results suggest that, in addition to abstracting the relevant dimensions, *S*s learn any given classification by formulating a rule for building that classification up out of Type I classifications. This tentative characterization of the learning process seems to provide a basis for understanding the obtained ranking of the six types of classifications with respect to initial difficulty, the markedly greater positive transfer in Type VI classifications, the effect upon difficulty of using compact or distributed stimuli, and certain individual differences.

REFERENCES

Archer, E. J., Bourne, L. E., Jr., & Brown, F. G. Concept identification as a function of irrelevant information and instructions. *J. exp. Psychol.*, 1955, **49**, 153–164.

Baum, Marian H. Simple concept learning as a function of intralist generalization. *J. exp. Psychol.*, 1954, **47**, 89–94.

Binder, A., & Feldman, S. E. The effects of experimentally controlled experience upon recognition responses. *Psychol. Monogr.*, 1960, **74**(9, Whole No. 496).

Bourne, L. E., Jr., & Haygood, R. C. The role of stimulus redundancy in concept identification. *J. exp. Psychol.*, 1959, **58**, 232–238.

Bourne, L. E., Jr., & Restle, F. Mathematical theory of concept identification. *Psychol. Rev.*, 1959, **66**, 278–296.

Bricker, P. D. The identification of redundant stimulus patterns. *J. exp. Psychol.*, 1955, **49**, 73–81.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. *A study of thinking.* New York: Wiley, 1956.

Bush, R. R., & Mosteller, F. A model for stimulus generalization and discrimination. *Psychol. Rev.*, 1951, **58**, 413–423.

Buss, A. H. A study of concept formation as a function of reinforcement and stimulus generalization. *J. exp. Psychol.*, 1950, **40**, 494–503.

Dattman, Priscilla E., & Israel, H. E. The order of dominance among conceptual capacities: An experimental test of Heidbreder's hypothesis. *J. Psychol.*, 1951, **31**, 147–160.

Estes, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.

Estes, W. K. Of models and men. *Amer. Psychologist*, 1957, **12**, 609–617.

Estes, W. K. Component and pattern models with Markovian interpretations. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford, Calif.: Stanford Univer. Press, 1959. Pp. 9–52. (a)

Estes, W. K. The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: A study of a science.* Vol. 2. New York: McGraw-Hill, 1959. Pp. 383–491. (b)

Estes, W. K. Statistical models for recall and recognition of stimulus patterns by human observers. In M. C. Yovits & S. Cameron (Eds.), *Self-organizing systems.* New York: Pergamon, 1960. Pp. 51–62.

French, R. S. A comparison of rote and discrimination learning with stimulus homogeneity variant. Unpublished doctoral dissertation, University of Wisconsin, 1951.

French, R. S. Number of common elements and consistency of reinforcement in a discrimination learning task. *J. exp. Psychol.*, 1953, **45**, 25–34.

Gibson, Eleanor J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.

Heidbreder, Edna. The attainment of concepts: II. The problem. *J. gen. Psychol.*, 1946, **35**, 191–223.

Heidbreder, Edna. The attainment of concepts: III. The process. *J. Psychol.*, 1947, **24**, 93–138.

Higonnet, R. A., & Grea, R. A. *Logical design of electrical circuits.* New York: McGraw-Hill, 1958.

Hovland, C. I. A "communication analysis" of concept learning. *Psychol. Rev.*, 1952, **59**, 461–472.

Hovland, C. I., & Hunt, E. B. Computer simulation of concept attainment. *Behav. Sci.*, 1960, **5**, 265–267.

Hovland, C. I., & Weiss, W. Transmission of information concerning concepts through positive and negative instances. *J. exp. Psychol.*, 1953, **45**, 175–182.

Hull, C. L. Quantitative aspects of the evolution of concepts. *Psychol. Monogr.*, 1920, **28**(1, Whole No. 123).

Keller, F. S., & Taubman, R. E. Studies in international Morse code: II. Errors made in code reception. *J. appl. Psychol.*, 1943, **27**, 504–509.

McGill, W. J. Multivariate information transmission. *Psychometrika*, 1954, **19**, 97–116.

Metzger, R. A comparison between rote learning and concept formation. *J. exp. Psychol.*, 1958, **56**, 226–231.

Morrisett, L., Jr., & Hovland, C. I. A comparison of three varieties of training in human problem solving. *J. exp. Psychol.*, 1959, **58**, 52–55.

Newell, A., Shaw, J. C., & Simon, H. A. Empirical explorations of the logic theory machine. In, Proceedings of the western joint computer conference, Institute of Radio Engineers, 1957.

Newell, A., & Simon, H. A. The simulation of human thought. (RAND Paper P-1734) Santa Monica, Calif.: RAND, 1959.

Newman, S. E. Effects of contiguity and similarity on the learning of concepts. *J. exp. Psychol.*, 1956, **52**, 349–353.

Oseas, L., & Underwood, B. J. Studies of distributed practice: V. Learning and retention of concepts. *J. exp. Psychol.*, 1952, **43**, 143–148.

Plotkin, L. Stimulus generalization in Morse code learning. *Arch. Psychol., NY*, 1943, **40**, No. 287.

Restle, F. A theory of discrimination learning. *Psychol. Rev.*, 1955, **62**, 11–19.

Restle, F. Theory of selective learning with probable reinforcements. *Psychol. Rev.*, 1957, **64**, 182–191.

Rogers, M. S. An application of information theory to the problem of the relationship between meaningfulness of material and performance in a learning situation. Unpublished doctoral dissertation, Princeton University, 1952.

Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.*, 1958, **65**, 386–408.

Shepard, R. N. Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, 1957, **22**, 325–345.

Shepard, R. N. Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychol. Rev.*, 1958, **65**, 242–256. (a)

Shepard, R. N. Stimulus and response generalization: Tests of a model relating generalization to distance in psychological space. *J. exp. Psychol.*, 1958, **55**, 509–523. (b)

Smith, S. L. Concept formation: Stimulus dimensions in human learning. Unpublished doctoral dissertation, Massachusetts Institute of Technology, 1954.

Sutcliffe, J. P. A general method of analysis of frequency data for multiple classification designs. *Psychol. Bull.*, 1957, **54**, 134–137.

Wallach, Lise. The complexity of concept attainment. *Amer. J. Psychol.*, in press.

Wolfle, D. L. The relation between linguistic structure and associative interference in artificial linguistic material. *Lang. Monogr.*, 1932, No. 11.

(Received October 25, 1960)

APPENDIX

The purpose of this section is to show how certain information-theoretic considerations lead in a rather natural way to the obtained ranking of the six types of classifications with respect to initial difficulty. The variables x, y, and z will be used to represent the three dimensions of the stimuli: e.g., size, color, and shape. Thus the variable x might take on the two values: large and small. In addition, a classification variable, C, can be defined to take on one value (e.g., A) for all stimuli in one class and another (B) for all stimuli in the other class. For all the classifications considered here, our initial uncertainty about the value of the classification variable for a randomly selected stimulus is just one bit. If we know exactly what classification is involved, a complete specification of the stimulus (i.e., a statement of its size, color, and shape) will reduce our uncertainty about the value of the classification variable from one bit to zero. This total reduction in uncertainty can be partitioned into three components: $C(x)$, the reduction due to the specification of size alone; $C_x(y)$, the reduction due to the additional specification of color (over and above that due to the specification of size alone); and $C_{xy}(z)$, the reduction due to the additional and final specification of shape (over and above that due to the specification of size and color together). Since the total reduction is one bit:

$$C(x) + C_x(y) + C_{xy}(z) = 1$$

This equation shows, in one way, the extent to which the information about the classification variable is spread out among the three dimensions of the stimuli. But it is arbitrary in that the three variables x, y, and z could be taken in any of their 3! or 6 possible orders. And, except for classifications of Types IV and VI, the magnitudes of the three terms will depend upon the order in which the variables are taken. However, we shall suppose that the variables are taken in their optimum order: i.e., the order for which the one-variable term is as large as possible and the three-variable term is as small as possible. Assuming, then, that the variables are taken in this order, we shall rewrite the equation in the simpler form:

$$C_1 + C_2 + C_3 = 1$$

The magnitudes of the one-, two-, and three-variable terms can readily be calculated for each of the six types of classifications (see McGill, 1954). They are 1,0,0 for Type I; 0,1,0 for Type II; 0,0,1 for Type VI; and 0.189, 0.311, 0.500 for each of the three remaining types, III, IV, and V.

Now the difficulty of extracting information presumably increases with the number of variables to which *S*s must simultaneously attend in order to extract that information. Therefore the difficulty of a classification should be greater if a larger fraction of the information about the classification is contained in the two-variable and, particularly, the three-variable terms. An index of difficulty, D, could therefore be defined by weighting each term according to the number of variables involved in that term. This can be done in a general way by writing:

$$C_1 + \alpha C_2 + \beta C_3 = D$$

If $\beta = \alpha = 1$, this reduces to the previous equation and the resulting index is unity for all six types of classifications. But, if $\beta > \alpha > 1$, the resulting index ranks the six types—presumably with respect to difficulty.

For any particular choice of values for α and β we can determine the ranking implied by those values by substituting them (together with the numerical values already determined for C_1, C_2, and C_3) into the expression for the difficulty, D. Figure A1 shows the result of a systematic exploration of the rankings implied by every pair of coefficients, α and β, for which $\beta > \alpha > 1$. As can be seen, only three solutions are possible: if $\beta < 1.378\,\alpha - 0.378$, then I < (III, IV, V) < II < VI; if $\beta = 1.378\,\alpha - 0.378$, then I < (II, III, IV, V) < VI; and if $\beta > 1.378\,\alpha - 0.378$, then I < II < (III, IV, V) < VI. Moreover, since the number of possible configurations of values along the dimensions of the stimuli increases exponentially with the number of dimensions, the increment in difficulty that results from attending to three dimensions instead of two is probably at least as great as the increment that results from attending to two dimension instead of one. Therefore we should expect that $(\beta - \alpha) \geqq (\alpha - 1)$ and, a fortiori, that $\beta > 0.378\,\alpha - 0.378$. Hence, we again arrive at the ranking I < II < (III, IV, V) < VI.

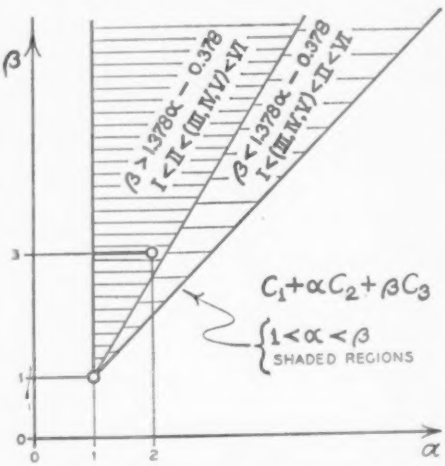

FIG. A1. Space of possible coefficients, α and β, for the measure, D, of the difficulty of a classification. (The indicated point for $\alpha = 2$ and $\beta = 3$ corresponds to the case in which each term in the expression for D is weighted by the number of variables involved in that term. This point might be considered to represent a reasonable choice of values for α and β.)